(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 November 2001 (22.11.2001)

PCT

(10) International Publication Number
WO 01/88188 A2

(51) International Patent Classification[7]: C12Q 1/68, G06F 19/00

(21) International Application Number: PCT/JP01/04192

(22) International Filing Date: 18 May 2001 (18.05.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2000-145977 18 May 2000 (18.05.2000) JP

(71) Applicant *(for all designated States except US)*: **NIHON UNIVERSITY, SCHOOL JURIDICAL PERSON** [JP/JP]; 8-24, Kudan-minami 4-chome, Chiyoda-ku, Tokyo 102-8275 (JP).

(72) Inventors; and

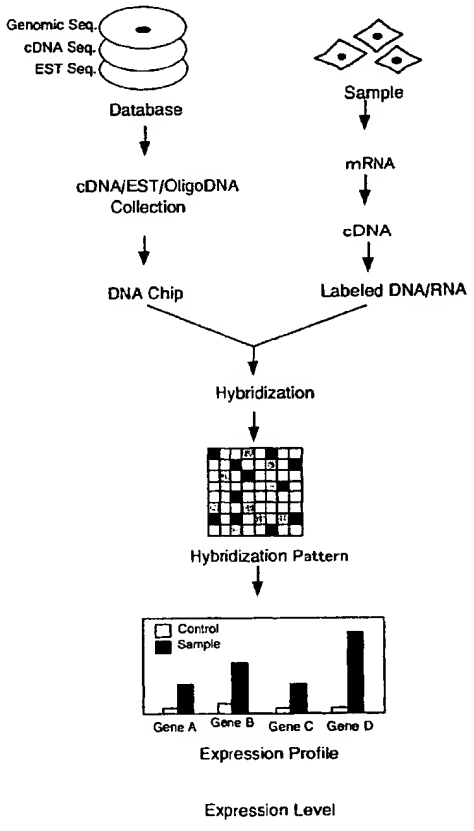
(75) Inventors/Applicants *(for US only)*: **ISHIKAWA, Koichi** [JP/JP]; 28-32-1304, Hongo 1-chome, Bunkyo-ku, Tokyo 113-0033 (JP). **ASAI, Satoshi** [JP/JP]; 36-10-1106, Higashi-ohizumi 1-chome, Nerima-ku, Tokyo 178-0063 (JP). **TAKAHASHI, Yasuo** [JP/JP]; 3-8-405, Mukaihara 2-chome, Itabashi-ku, Tokyo 173-0036 (JP). **NAGATA, Toshihito** [JP/JP]; 52-2-205, Ohyaguchi-kamicho, Itabashi-ku, Tokyo 173-0032 (JP). **ISHII, Yukimoto** [JP/JP]; 22-10, Sekibara 2-chome, Adachi-ku, Tokyo 123-0852 (JP).

(74) Agents: **HIRAKI, Yusuke** et al.; Toranomon No.5 Mori Building Third Floor, 17-1, Toranomon 1-chome, Minato-ku, Tokyo 105-0001 (JP).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK,

[Continued on next page]

(54) Title: METHOD FOR EXAMINING ISCHEMIC CONDITIONS



(57) Abstract: The present invention provides a method for examining ischemic conditions, comprising measuring the expression levels of particular genes in a test sample or determining the expression profile of a gene group in the sample comprising a plurality of genes selected from said particular genes.

WO 01/88188 A2

SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

DESCRIPTION

METHOD FOR EXAMINING ISCHEMIC CONDITIONS

TEHCHNICAL FIELD

The present invention relates to a method for examining ischemic conditions by measuring the expression levels of particular genes in a test sample or by determining the expression profile of a gene group in the sample comprising a plurality of genes selected from the particular genes.

BACKGROUND ART

Cancer, cerebral apoplexy and heart diseases are called three major adult diseases and they occupy about 60% of the causes of death in the Japanese. Of these, cerebral apoplexy and heart diseases are often caused by ischemia. Thus, early detection of ischemic conditions makes it possible to prevent these diseases from occurring. Ischemia is local anemia and may be classified into groups such as compressive ischemia caused by constriction or occlusion in arterial walls due to external pressure from tumor or the like; occlusive ischemia caused by changes inside the blood vessels or in the blood vessels themselves such as thrombosis or arterial sclerosis; and vasospastic ischemia caused by vasospasms such as cerebral anemia or angina, from the viewpoint of the mechanism of its occurrence. Ischemia in the brain triggers ischemic cerebral apoplexy such as cerebral infarction, and ischemia in the heart triggers ischemic heart diseases such as myocardial infarction. Thus, for the prevention of these diseases, it is important, first of all, to find ischemic conditions as early as possible and to receive appropriate treatment.

As a method for examining ischemic conditions, a method in which abnormality in the cardiac wall movement is used as an indicator (e.g., quantitative analysis of ventricular forms/ultrasonic images, or detection of decrease in tissue systole speed); a method in which abnormality in hemodynamics is used as an indicator (e.g., analysis of the pattern of blood flow rate into the left ventricle, or nuclear medicine examination) and the like have been known to date. Among all, nuclear medicine examination is a method which can

examine ischemic conditions accurately. However, this method is disadvantageous to subjects because it involves exposure to radiation, requires a long time for examination, and is expensive. Under circumstances, a simple examination method for ischemic conditions has been desired which imposes less burden to subjects and can be carried out routinely as a part of health examination.

DISCLOSURE OF THE INVENTION

It is an object of the present invention to provide a method for examining ischemic conditions simply by measuring the expression levels of particular genes in a test sample or by determining the expression profile of a gene group in the sample comprising a plurality of genes selected from the particular genes.

As a result of intensive and extensive researches toward the solution of the above problems, the present inventors have succeeded in identifying those genes expressed under ischemic conditions and in elucidating the expression profile of the genes. Thus, the present invention has been achieved.

The present invention relates to a method for examining ischemic conditions, comprising measuring the expression levels of particular genes in a test sample or determining the expression profile of a gene group in the sample comprising a plurality of genes selected from the particular genes. Specific examples of the particular genes include (a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or (b) genes functionally equal to the genes having any of the above-described nucleotide sequences or genes functionally equal to the genes encoding any of the above-described amino acid sequences. The measurement of expression levels and the determination of expression profile may be carried out using a DNA chip (e.g., a synthetic-type DNA chip). Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

Further, the present invention relates to a DNA chip for examining ischemic conditions, which carries a part or all of the following genes (a) or (b) immobilized on its surface: (a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through

SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or (b) genes functionally equal to the genes having any of the above-described nucleotide sequences or genes functionally equal to the genes encoding any of the above-described amino acid sequences. Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

Further, the present invention relates to a method of screening for ischemic condition-improving drugs or therapeutics for ischemic diseases. This method is characterized by selecting candidate drugs using as an indicator whether or not:

(a) the expression levels of particular genes of which expression levels change under ischemic conditions return to a normal expression levels; or

(b) the expression profile of a gene group comprising a plurality of the particular genes returns to a normal expression profile;

by the administration of a drug to a test animal or test cell, wherein the returning to the normal expression levels or normal expression profile indicates that the drug is a candidate drug. Specific examples of the particular genes include (a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or (b) genes functionally equal to the genes having any of the above-described nucleotide sequences or genes functionally equal to the genes encoding any of the above-described amino acid sequences. Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

Further, the present invention relates to a computer-readable record medium in which the following data (i) or (ii) have been recorded: (i) expression level data of particular genes of which expression levels change under ischemic conditions, or (ii) expression profile data of a gene group comprising a plurality of genes selected from the particular genes. Specific examples of the particular genes include (a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or (b) genes functionally equal to the genes having any of the above-described

nucleotide sequences or genes functionally equal to the genes encoding any of the above-described amino acid sequences. Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

Further, the present invention relates to a computer-readable record medium in which a program that directs a computer to execute the following procedures has been recorded:

(a) procedures to input expression level data or expression profile data of particular genes in a test sample;

(b) procedures to record the input data;

(c) procedures to check the recorded data with already recorded expression level data or expression profile data of the particular genes under ischemic conditions;

(d) procedures to determine whether the test sample is under ischemic conditions or not based on the checking results obtained in (c); and

(e) if the test sample has been determined as being under ischemic conditions, procedures to identify the clinical stage of the ischemic conditions of the test sample based on the checking results obtained in (c). Specific examples of the particular genes include (a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or (b) genes functionally equal to the genes having any of the above-described nucleotide sequences or genes functionally equal to the genes encoding any of the above-described amino acid sequences. Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

Hereinbelow, the present invention will be described in detail.

The present specification encompasses the contents of the specification and drawings of Japanese Patent Application No. 2000-145977 based on which the present application claims priority.

The present invention relates to a unique method for examining ischemic conditions using, as an indicator, expression levels of particular genes or an expression profile of a particular gene group. The term "expression level" used herein refers to an absolute or

relative amount of the transcript (i.e., mRNA) of a particular gene; or an absolute or relative amount of the translation product (i.e., protein) of a particular gene. The term "expression profile" used herein refers to expression levels of a plurality of genes collected and arranged in tables, graphs, or the like.

1. Identification of Genes of which Expression Levels Change under Ischemic Conditions

Genes of which expression levels change under ischemic conditions may be identified by, for example, the differential RNA display method [Liang, P. et al., Science 257:967-971 (1992)], the hybrid subtraction method, or a method using a DNA chip. For example, a method of identifying the above genes using a DNA chip may be carried out as illustrated in Fig. 1. Briefly, a plurality of pieces of DNA information (cDNA/EST/oligoDNA collection) are obtained from a DNA database where genomic sequences, cDNA sequences or EST sequences have been recorded. Then, a wide variety of genes of known sequences are immobilized on a DNA chip. Subsequently, labeled DNA or RNA which has been prepared from mRNA derived from biosamples under ischemic conditions is hybridized with the DNA chip. The hybridization strength at each spot of the resultant hybridization pattern is then measured to thereby measure the expression level of each gene. Thus, an expression profile is obtained. Hereinbelow, a method using a DNA chip will be described in more detail.

(1) Preparation of Poly(A)+mRNA from Test Samples

First, poly(A)+mRNA must be prepared from test samples such as tissue or cell to examine expression levels of particular genes in the samples using a DNA chip. Specific examples of test samples useful for the preparation of poly(A)+mRNA to be used in the identification of genes of which expression levels change under ischemic conditions include biotissues (e.g., blood tissue, brain tissue, heart tissue or renal tissue) derived from experiment animals (e.g., mice, rats, guinea pigs, rabbits, dogs, cats, pigs or cows) in which ischemic conditions have been induced artificially or derived from humans under ischemic conditions. It is said that about 80% of genes which may be expressed in a living body are being expressed in the brain. Thus, by examining those genes of which expression levels

change in the brain under ischemic conditions, it is possible to comprehensively identify those genes of which expression levels change under ischemic conditions in tissues other than the brain. More specifically, the hippocampus derived from the above-described mice may be used as a test sample for preparing poly(A)+mRNA. Since cerebral capillaries exist in the hippocampus, blood cells such as erythrocytes, leukocytes and platelets are present there in a mixed state. Therefore, mRNAs from various blood cells may be contained in mixture in poly(A)+mRNA extracted from a hippocampus tissue.

Poly(A)+mRNA may be prepared by obtaining total RNA from test samples by such methods as the guanidine thiocyanate-cesium chloride method [J. Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, New York (1989)], the guanidine thiocyanate-cesium trifluoroacetate method [H. Okayama et al., Methods in Enzymology, 154:3, Academic Press, New York (1987)], the guanidine thiocyanate-phenol-chloroform method [P. Chomczynski et al., Anal. Biochem., 162:156 (1987)] or the phenol-SDS method [R.D. Palmiter, Biochemistry, 13:3606 (1974)], loading the resultant total RNA to an oligo-dT cellulose or poly U Sepharose column for specific adsorption of poly(A)+mRNA, and then eluting the poly(A)+mRNA from the column. In particular, when the test sample is a tissue, it is important to perform purification processes accurately because the state of purification of the total RNA or poly(A)+mRNA greatly influences the yield of cDNA, etc.

For example, when poly(A)+mRNA is prepared by the guanidine thiocyanate-cesium chloride method, first, an appropriate amount (e.g., 5 volumes) of guanidine thiocyanate solution is added to a tissue sample. Then, the tissue sample is disrupted using, e.g., a Polytron homogenizer. Sodium N-lauroyl sarcosinate is added to the disrupted tissue to give a desired concentration (e.g., 0.5%) and agitated. The resultant sample is centrifuged (e.g., at 5000 xg for 10 min). The resultant supernatant is layered over a cushion of cesium chloride-EDTA contained in a centrifuge tube and subjected to ultracentrifugation (e.g., at 100,000 xg for 12 hr). The resultant precipitate is rinsed with 70% ethanol and then dissolved in TE buffer to thereby obtain total RNA. The resultant total RNA is applied to an oligo-dT cellulose column to thereby obtain poly(A)+mRNA.

Alternatively, the preparation of poly(A)+mRNA may be performed using

commercial kits. Specific examples of kits for preparing total RNA include RNeasy Total RNA Isolation kit (Qiagen) and TRIzol Reagent (Gibco BRL Life Technologies). Specific examples of kits for isolating poly(A)+mRNA from total RNA include Oligotex Direct mRNA kit (Qiagen) and Oligotex mRNA kit (Quiagen).

(2) Synthesis of cDNA with Reverse Transcriptase

Subsequently, cDNA is synthesized using the poly(A)+mRNA obtained in (1) above as a template. The synthesis of cDNA may be carried out according to the method of Gubler et al. [U. Gubler et al., Gene 25:263 (1987)]. Briefly, oligo(dT)$_{12-18}$ is added to a solution of poly(A)+mRNA, which is heated and then cooled quickly. To this solution, a single-stranded cDNA synthesis buffer, a dNTP solution (containing mixture of dATP, dGTP, dCTP and dTTP), a ribonuclease inhibitor solution, a dithiothreitol solution, etc. are added and mixed. Then, a reverse transcriptase (e.g., Superscript RT; BRL) is added to the mixture, which is then incubated for a specific period to thereby yield single-stranded cDNA. If necessary, double-stranded cDNA may be synthesized further using the single-stranded cDNA as a template. Briefly, a cDNA synthesis buffer, a dNTP solution (containing mixture of dATP, dGTP, dCTP and dTTP), a dithiothreitol solution, etc. are added to a solution of the single-stranded cDNA and mixed. Then, a DNA polymerase (e.g., T4 DNA polymerase) is added to the mixture, which is then incubated for a specific period to thereby yield double-stranded cDNA. Labeled cDNA may be obtained by using a labeled dNTP (e.g., biotin-labeled dNTP) in the synthesis of single- or double-stranded cDNA.

(3) Preparation of Labeled cRNA Fragments

When a DNA chip on which oligonucleotides are immobilized as DNA probes is used in the method of the invention, labeled cRNA is prepared, if necessary, by *in vitro* transcription using the cDNA obtained in (2) above as a template. The preparation of labeled cRNA by *in vitro* transcription may be carried out according to the method of Kreig et al. [Kreig, P.A. et al., Methods in Enzymology 155:397-415 (1987)]. The resultant labeled cRNA molecules must be fragmented before use. The fragmentation of these molecules may be performed by heating in the presence of Mg^{2+} (e.g., at 94℃ for 3 min) or by treatment with DNase.

The *in vitro* transcription described above may also be performed using a commercial kit. As an example of *in vitro* transcription kit, MEGAscript™ *In Vitro* Transcription Kit (Ambion) may be given.

(4) Hybridization on a DNA Chip

Subsequently, the labeled nucleotide sample obtained in (2) or (3) above is added to a DNA chip to carry out a hybridization reaction. Specific examples of DNA chips useful in the method of the invention include oligoDNA microarray (also called "synthetic-type DNA chip") which is prepared by synthesizing oligoDNAs on a substrate directly, and DNA microarray (also called "paste-type DNA chip") which is prepared by immobilizing pre-synthesized DNAs on a substrate. In the present invention, it is preferable to use a synthetic-type DNA chip that can provide high detection sensitivity, accuracy and reproducibility (e.g., oligoDNA microarray GeneChip™ manufactured by Affymetrix) for identifying genes of which expression levels change under ischemic conditions.

In the examination of gene expression, it is important to carry out hybridization under high stringency conditions to inhibit non-specific bonding. The term "high stringency conditions" refers to those conditions under which hybridization only occurs between two nucleotide strands having 90% or more homology to each other. Stringency may be raised or lowered by changing salt concentrations (e.g., concentrations of NaCl, trisodium citrate) and/or the reaction temperature. The lower the salt concentrations are and the higher the temperatures is, the higher the stringency becomes. Depending on the type of DNA chip used and other factors, a specific temperature and specific salt conditions may be high stringency conditions or low stringency conditions. Thus, high stringency conditions and low stringency conditions should be decided for each chip to be used. With respect to GeneChip™ Mu6500 used in the present invention, high stringency conditions refer to reaction temperatures ranging from 43 to 65℃, preferably 45℃, and Na^+ concentrations ranging from 500 to 1000 mM, preferably 1000 mM.

(5) Detection and Data Analysis

The double-strands formed on the microarray as a result of the hybridization are analyzed with a fluorescence image scanner or the like. The fluorescence intensities may be measured automatically with a system integrating a fluorescent laser microscope, a CCD

camera and a computer. Preferably, a scanner is used which is capable of quantitatively discriminating spots having a size of several ten micrometers and having a distance of approx. 10 μm between every two spots. Further, it is preferable that the scanner be capable of handling a plurality of labels and scanning over a wide range at a high speed, and that the scanner be equipped with an automatic focusing function which allows the scanner to manage microscopic distortion in the substrate. As a specific example of a scanner equipped with such a function, GMS 418 Array Reader (Genetic MicroSystems) may be given. The software to be used for the analysis of the above data is, preferably, capable of performing complicated analysis of a large number of oligonucleotides with partially overlapped sequences, such as analysis of mutation or polymorphism.

Alternatively, a commercial system may be used in the present invention which is integrating a set of components necessary for gene analysis using a DNA chip. These components include (i) a DNA chip, (ii) a device for automatically washing and staining the DNA chip after hybridization, (iii) a scanner which reads fluorescence emission, and (iv) a work station which processes and analyzes the information read. As a specific example of such a system, the GeneChip™ analysis system created by Affimetrix may be given. This system is provided with GeneChip™ Laboratory Information Management System (LIMS™) and GeneChip™ Expression Data Mining Tool (EDMT™) as bioinformatics tools for efficient utilization of obtained genetic data. These tools make it possible to output obtained data to SQL compliant databases of GATC (Genetic Analysis Technology Consortium) format to thereby link the system to public genetic information databases on the web. By using this analysis system, more efficient and more extensive data analysis can be made.

The nucleotide sequences of those genes which have been found by the present invention to show altered expression levels under ischemic conditions in mouse, and the amino acid sequences encoded by those genes are shown in SEQ ID NO: 1 through SEQ ID NO: 1066. Since both mouse and human belong to mammals, they are highly similar to each other genetically. Thus, genes which are functionally equivalent to the genes having the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or the genes encoding the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066

may exist in human cells. Accordingly, by measuring expression levels of such human genes, it is possible to perform examination of ischemic conditions on human-derived samples. The term "functionally equivalent genes" used herein includes, in addition to the genes consisting of any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 themselves or the genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 themselves, those genes which have homology to the above genes and play roles identical or similar to the roles of the above genes in the living body. Nucleotide sequence information, amino acid sequence information, etc. on those genes in human cells which are functionally equivalent to mouse-derived genes can be obtained from known databases such as GenBank by searching with keywords such as a part of nucleotide sequence of interest, a part of amino acid sequence of interest, or a gene product name.

It is possible to identify ischemia marker genes of which expression levels change specifically under ischemic conditions. This identification can be performed by further examining the expression levels under other diseases of the above-described genes which were found to show altered expression levels under ischemic conditions.

2. DNA Chips for Examining Ischemic Conditions

DNA chips carrying as DNA probes a part or all of the genes identified in Section 1. above (which show altered expression levels under ischemic conditions) immobilized on their substrates can be used as a DNA chip for examining ischemic conditions. In particular, a DNA chip as shown in Fig. 2 on which three groups of genes (i.e., genes showing high expression levels under ischemic conditions; genes showing moderate expression levels under ischemic conditions; and genes showing low expression levels under ischemic conditions) are located separately may be used as a DNA chip that is capable of evaluating the extent of progress of ischemic conditions. There are two types of DNA chips. One is paste-type chips prepared by immobilizing pre-synthesized DNA probes on their substrates; and the other is synthetic-type chips prepared by synthesizing DNA probes on their substrates directly. The term "DNA probes" used herein refers to DNA strands which are immobilized on the substrate of a DNA chip in order to detect those

genes having DNA strands with specific nucleotide sequences. Hereinbelow, processes for preparing both types of DNA chips will be described specifically.

(1) Method of Preparing Paste-Type DNA Chips

First, as DNA probes, a part or all of the genes identified in Section 1. above which show altered expression levels under ischemic conditions are prepared by PCR or chemical synthesis. DNA probes must be present on the substrate of a DNA chip as single-stranded DNAs so that they can hybridize with target nucleotide strands having sequences complementary to the sequences of the DNA probes when the target strands access to the DNA probes. Thus, in designing DNA probes, it is desirable to select sequences so that formation of secondary structures that would inhibit the hybridization with target nucleotide strands will occur as little as possible. The term "secondary structures" used herein refers to the stem-loop structure, hairpin structure or the like which is formed by hybridization of a part of a probe with another part of the same probe when the probe has been folded back. Whether sequences of interest would form secondary structures or not can be analyzed using a commercial gene analysis software (e.g., DNASIS; Hitachi Software Engineering).

The preparation of DNA probes by PCR may be carried out by conventional methods [see, for example, Sambrook, J. et al., Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press (1989)] using, as a template, genomic DNA, total RNA, mRNA or cDNA derived from an organism to be tested. For example, the gene consisting of the nucleotide sequence shown in SEQ ID NO: 1006 can be used as a marker for examining ischemic conditions since its expression increases remarkably under ischemic conditions. A DNA probe to detect this gene may be obtained by PCR using sense primer 5'-atgctcttccgagctgtgct-3' (SEQ ID NO: 1067), anti-sense primer 5'-cagctcagttgaacgccttt-3' (SEQ ID NO: 1068) and, as a template, cDNA prepared from mRNA derived from mouse hippocampus under ischemic conditions. Whether the amplified fragment by PCR is the fragment of interest or not may be determined by subcloning the amplified fragment into an appropriate vector such as pBlueScriptSK(+) (Stratagene) or pCR2.1 (Invitrogen) and then determining the nucleotide sequence thereof. The nucleotide sequence may be determined by conventional methods such as the chemical modification method by Maxam-Gilbert or

the dideoxynucleotide chain termination method using M13 phage. Usually, the nucleotide sequence may be determined using an automated DNA sequencer (e.g., 373A DNA sequencer; Perkin-Elmer).

On the other hand, the preparation of DNA probes by chemical synthesis may be carried out according to conventional DNA synthesis methods used in the art, e.g., the phosphoramidite method, or the phosphonate method. For example, when DNA probes are synthesized by the phosphoramidite method, a nucleoside derivative obtained by introducing a trivalent phosphoramidite residue into the hydroxyl group at 3'-position of the sugar moiety is used as a synthesis unit. Fist, this amidite unit is activated with 1H-tetrazol and reacted with the 5'-hydroxyl of a DNA strand on a solid phase (step 1), to thereby yield a trivalent phosphite ester. Subsequently, the trivalent phosphite ester is led to a pentavalent phosphate triester through oxidation (step 2), capping (step 3) and hydrogenation (step 4). Then, steps 1 to 4 are repeated. Finally, an oligomer block having the desired nucleotide sequence is cleaved from the solid phase and deprotected to thereby yield the DNA strand of interest.

Subsequently, the thus obtained DNA probe is immobilized on the substrate of a DNA chip. Specific examples of substrates useful for this purpose include glass sheets, quartz sheets and silicone wafers. As a size of the substrate, 3.5 mm x 5.5 mm, 18 mm x 18 mm or 22 mm x 75 mm may be used, for example. This size may be varied appropriately depending on, for example, the number and size of spots of DNA probes on the substrate. As to a method for immobilizing DNA, DNA may be electrostatically bound to a solid support that has been surface-treated with a polycation such as polylysine, polyethyleneimine or polyalkylamine, utilizing the electric charge of the DNA; or DNA probes into which a functional group such as amino group, aldehyde group, SH-group or biotin has been introduced may be covalently bound to the surface of a solid support into which a functional group such as amino group, aldehyde group or epoxy group has been introduced.

The spotting of DNA probes on the substrate may be performed using an arrayer which is capable of quantitatively spotting DNA probes in sizes ranging from several ten micrometers to several hundred micrometers and at pre-determined locations. As to the

technology of spotting, pin technology utilizing the mechanical contact of pin tips with a solid support; inkjet technology utilizing the principle of inkjet printer; or capillary technology utilizing a capillary device may be enumerated.

(2) Method for Preparing Synthetic-Type DNA Chips

As a method for synthesizing DNA probes on a substrate directly, the method of Fodor et al. may be used in which photolithographic fabrication techniques are combined with solid phase DNA synthesis techniques [Foder, S.P.A. et al, Science 251:767-773 (1991)]. Briefly, a synthetic linker having a protective group removable by a photochemical reaction is bound onto a substrate. Then, the substrate is illuminated by light through a blocking material called mask to thereby remove only those protective groups in specific areas. Subsequently, the substrate is reacted with nucleotides having protected hydroxyl groups. As a result, polymerization occurs only in those areas where protective groups have been removed. Then, the substrate is illuminated by light through another mask, and polymerization of nucleotides is repeated. Thus, coupling reactions with different nucleotide precursors are repeated using various masks. As a result, DNA probes of desired sequences can be synthesized on specific areas on the substrate of a DNA chip. An oligonucleotide N-mer in nucleotide length can be synthesized by Nx4 cycles of reaction. Thus, a DNA probe 25-mer in length can be synthesized by 25 x 4 = 100 cycles of reaction. The nucleotide length of the DNA probes on the DNA chip of the invention for examining ischemic conditions is 10- to 30-mer, preferably 15- to 25-mer.

Since the nucleotide length of DNA probes on DNA chips of this type is usually short, the specificity of hybridization on such chips may be questioned. This problem can be solved as described below. Briefly, in order to detect the expression of a particular gene, perfect match (PM) (i.e., completely complementary) oligonucleotide DNA probes corresponding to a plurality of portions (usually, ten and several portions) of the target gene and an identical number of mismatch (MM) oligonucleotide DNA probes having a mutation at one nucleotide (usually, the central nucleotide or neighboring nucleotide) are located on a substrate (see Fig. 3). Then, hybridization is carried out on the substrate using the MM probes as an indicator of the specificity of hybridization. That is, signal ratio of PM probes to MM probes is calculated, and the pseudo-positive signal is

eliminated.

3. Method of the Invention for Examining Ischemic Conditions

Ischemic conditions can be examined by measuring the expression levels in a test sample of the genes which were revealed in the present invention to show altered expression levels under ischemic conditions. Alternatively, ischemic conditions can be examined by determining the expression profile of a gene group comprising a plurality of genes selected from the above-described genes. The expression "determining the expression profile of a gene group" means measuring the expression levels of individual genes constituting the group and arranging the results in tables, graphs, or the like. Specific examples of the ischemic conditions include compressive ischemia, occlusive ischemia and vasospastic ischemia.

By measuring the expression levels in a test sample of the genes identified in Section 1. above which show altered expression levels under ischemic conditions, it is possible to examine whether the test sample is under ischemic conditions or not. Briefly, when the expression levels of the above genes in the test sample are changed to the same extent as the expression levels under ischemic conditions revealed by the present invention are changed, then the test sample can be evaluated as being under ischemic conditions. For example, the expression levels of the genes having the nucleotide sequences shown in SEQ ID NOS: 960-1037 and SEQ ID NOS: 1065-1066 or the genes encoding the amino acid sequences shown in SEQ ID NOS: 960-1037 and SEQ ID NOS: 1065-1066 increase more than 10-fold under ischemic conditions. Thus, when the expression levels of these genes are increased more than 10-fold in a test sample, the test sample can be evaluated as being under ischemic conditions.

Further, it is possible to more accurately examine whether a test sample is under ischemic conditions or not by measuring the expression levels of ischemia marker genes, which are genes included in the genes identified in Section 1. above and show little or no changes in expression levels in diseases other than ischemia. That is, when the changes in the expression levels of the above marker genes in a test sample are the same in extent as the changes in the expression levels of those genes detected in the present invention, the test

sample can be evaluated as being under ischemic conditions.

The expression levels of the above genes may be measured by, for example, dot hybridization, slot hybridization, Northern hybridization or quantitative PCR when the number of genes to be measured is small. When the number of genes to be measured is large, their expression levels may be measured with a DNA chip.

Further, it is possible to examine with higher accuracy whether a test sample is under ischemic conditions or not by determining the expression profile of a gene group in the sample comprising a plurality of genes selected from the genes identified in Section 1. above which show altered expression levels under ischemic conditions and comparing the resultant profile with an expression profile of a normal sample which is not under ischemic conditions. The determination of expression profiles of gene groups can be performed more quickly and simply by using DNA chips. Expression profiles may be classified using cluster analysis described later. As to the DNA chip, it is preferable to use synthetic-type DNA chips from the viewpoints of accuracy, sensitivity and reproducibility. It is also possible to perform examination of ischemic conditions using the DNA chip of the invention prepared in Section 2. above. For example, as shown in Fig. 2A, genes are classified into a group of low expression level genes, a group of moderate expression level genes, and a group of high expression level genes and immobilized separately on a DNA chip for hybridization. Low expression level genes mean those genes of which transcription levels increased n-fold (where n is 2 or more but less than 5) within 24 hours when the transcription level at 0 hour is taken as 1. Moderate expression level genes mean those genes of which transcription levels increased n-fold (where n is 5 or more but less than 10). High expression level genes mean those genes of which transcription levels increased more than 10-fold.

In order to prepare a DNA chip carrying immobilized genes thereon which show altered expression levels under ischemic conditions, for example, 300 or more low expression level genes, 100-300 moderate expression level genes and 30-100 high expression level genes are selected from the genes having the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or the genes encoding the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066, and immobilized. When

the expression levels of the 30-100 high expression level genes have been changed compared with a gene expression profile obtained form a non-ischemic patient (control), the test sample can be judged under ischemic conditions of early-stage (Fig. 2B). When not only the expression levels of high expression level genes but also those of the 100-300 moderate expression level genes have been changed, the sample can be judged under ischemic conditions of intermediate-stage (Fig. 2C). Further, when the expression levels of the 300 or more low expression level genes have been changed in addition to those of the high expression level genes and the moderate expression level genes, the test sample can be evaluated under ischemic conditions of late-stage (Fig. 2D).

Expression levels may change toward increase or decrease compared to normal levels. However, the expression levels of all of the genes having the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or the genes encoding the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 change toward increase under ischemic conditions. Thus, in examining ischemic conditions using changes in the expression levels of the above genes as an indicator, a gene should be counted as having shown a change in its expression level only when an increase is observed in the expression level.

With respect to those genes which have been confirmed to show no change in expression levels under ischemic conditions, it is also possible to measure their expression levels and utilize the results for examining ischemic conditions. Briefly, the expression levels of such genes are measured in the test samples which were evaluated as ischemic as a result of the above-described measurement. If no change is observed in their expression levels, the reliability of the above evaluation that the test samples are under ischemic conditions can be enhanced.

4. Computer-Readable Record Medium Containing Gene Expression Data under Ischemic Conditions and Ischemic Condition Identification Program, as well as Ischemic Condition Identification Program

Expression level data of the genes which were revealed by the invention to show altered expression levels under ischemic conditions or expression profile data of a gene group comprising a plurality of genes selected from the above genes may be recorded in an

appropriate medium and used as comparison data in the analysis of examination data on ischemic conditions. As a record medium, any type of record media may be used, e.g., magnetic tape, CD-ROM, IC card, or RAM. Specifically, the degree of change in the expression levels of those genes of which expression levels change remarkably under ischemic conditions is examined in a test sample. If the expression levels are equal to the expression levels recorded in a record medium, the test sample (and thus the organism from which the sample is derived) can be evaluated as being under ischemic conditions. Further, more accurate evaluation of ischemic condition can be made by comparing a gene expression profile recorded (which was created from individual expression level data of a group of genes whose expression levels change under ischemic conditions) with an expression profile of corresponding genes in a test sample. That is, if the expression patterns in a test sample of a plurality of particular genes whose expression levels change under ischemic conditions resemble the expression patterns of the corresponding genes recorded in such a medium, the test sample can be evaluated as being under ischemic conditions with higher probability.

A record medium in which a program that directs a computer to execute the procedures described below has been recorded is useful as a record medium containing ischemic condition identification program. The term "ischemic condition identification program" used herein refers to a program that is able to identify the stage of ischemic conditions (i.e., early stage, intermediate stage or late stage) in a test sample when the test sample has been suspected to be under ischemic conditions or evaluated as being under such conditions. This program comprises (a) procedures to input expression level data or expression profile data of a test sample; (b) procedures to record the input data; (c) procedures to check this recorded data with already recorded expression level data or expression profile data under ischemic conditions; (d) procedures to determine whether the test sample is under ischemic conditions or not based on the checking results obtained in (c); and (e) if the test sample has been determined as being under ischemic conditions, procedures to identify the clinical stage of the ischemic conditions of the test sample based on the checking results obtained in (c). By analyzing the gene expression data of a test sample using a computer in which the above-described problem has been installed, ischemic

conditions can be identified.

The ischemic condition identification program of the invention comprises (a) means for analyzing expression levels of genes isolated from test cells; and (b) means for predicting whether or not individual test samples are under ischemic conditions or not, using the analysis results obtained by (a) as an indicator. The analysis means (a) is composed of means for detecting individual expression levels of a plurality of genes in test cells or tissues (also called "detection engine"), and means for analyzing the resultant values detected (also called "analysis engine").

(1) Detection Engine for Gene Expression

In the present invention, when expression of genes has been detected, the detection data may be digitalized and used as digital information. The digitalization is performed by converting, for example, fluorescence intensities detected on DNA chips into numerical values.

(2) Analysis Engine

Analysis engine is a means for performing analysis processing by multivariate analysis such as cluster analysis, based on the data (i.e., amounts of gene expression) obtained by the detection engine. Cluster analysis, which is a technique used in the field of multivariate analysis, collects and classifies "similar objects" from a large number of objects (i.e., samples) to be observed based on specific calculation criteria (assess criteria). In other words, cluster analysis merely "classifies" a large number of samples observed by putting samples similar to each other into one group.

In order to perform cluster analysis based on the detection data, "distance matrices" that represent similarities between samples are created. As the distance, Euclidean distance, weighted Euclidean distance, standard Euclidean distance, Pearson's product-moment correlation coefficient, or the like is calculated. The concepts of these distances are known, and an appropriate distance may be selected depending on the purpose of cluster analysis. Based on the concept of the above-mentioned distance, distances between clusters or distances between a cluster and objects are calculated, followed by amalgamation of clusters (i.e., two clusters are linked together). Methods of

amalgamation are known, e.g., the nearest neighbor method, furthest neighbor method, centroid method, or Ward's method.

By the above-described procedures, clusters which are in the "shortest distance" relation are linked together as "similar" clusters to thereby generate new clusters of a higher level. When clusters at one level have been generated, distances between clusters are calculated again to create distance matrices. Then, by searching for two clusters at the shortest distance, new clusters at a higher level are generated. Thus, a dendrogram is created finally.

Samples within a cluster amalgamated at a specific level of a dendrogram are contained in that cluster because of some similarity. Those samples with such similarity can be said to possess a certain nature in common. By elucidating this nature, it is possible to reveal the characteristic of the cluster itself. Thus, according to these analysis procedures, it is possible to classify genes into a group of high expression genes and a group of low expression genes. For example, if focusing on the degree of ischemic conditions using the progress (stage) of ischemia as an indicator, it is possible to reveal such characteristics that samples belonging to one cluster are under highly ischemic conditions and that samples belonging to the other cluster are under lightly ischemic conditions.

One embodiment of the identification system of the invention is illustrated in a block diagram (Fig. 5).

The identification system shown in Fig. 5 is equipped with CPU 501, ROM 502, RAM 503, Input Unit 504, Sending/Receiving Unit 505, Output Unit 506, Hard Disk Drive (HDD) 507 and CR-ROM Drive (508).

CPU 501 controls the ischemic condition identification system entirely and executes the examination processing described below according to the programs stored in ROM 502, RAM 503 or HDD 507. ROM 502 contains programs, etc. that instruct processing necessary for the operation of the above system. RAM 503 contains those data necessary for executing the examination processing. The Input Unit 504 is composed of a keyboard, mouse, etc. and operated, e.g., for inputting necessary conditions for the execution of the examination processing. The Sending/Receiving Unit 505 executes, based on instructions

from CPU 501, the sending/receiving of data to/from External Database 510, etc. through communication circuits. The Output Unit 506 executes display processing of various conditions input from the Input Unit 504, detection data on expressed genes, etc. based on instructions from CPU 501. The Output Unit 506 may include a computer-display unit and a printer. HD 507 contains information of expression patterns of various genes in cells or tissues and, based on instructions from CPU 501, reads out stored programs or data and stores them, e.g., in RAM 503. Based on instructions from CPU 501, CR-ROM Drive 508 reads out programs or data from the identification program stored in CD-ROM 509 and stores them, e.g., in RMA 503.

CPU 501 executes prediction of whether individual test samples are under ischemic conditions or not based on the data received from the Database, while supplying data received from the Input Unit, etc. to the Output Unit 506. The Database contains accumulated information of the amounts of gene expression (including both absolute amounts and relative amounts) obtained as described above.

Fig. 6 is a flow chart showing an example of identification processing using the identification program described above. Expression patterns of genes in test samples are analyzed, followed by identification of whether individual samples are under ischemic conditions or not .

Hereinbelow, the identification processing will be described with reference to Cluster Analysis Device 601 in Fig. 6. Cluster Analysis Device 601 generates clusters to be used in the identification processing. First, gene expression data are input by Means for External Database Searching and Data Input 602. Until data input is completed, input operation of the above data is repeated. By the input of the above data, information obtained from each tissue or cell is stored in Sample Data Storage Means 603, and supplied for cluster analysis or registered in the database.

Subsequently, Data Optimizing Means 604 inputs sample data from Sample Data Storage Means 603 and optimizes the data for cluster analysis. Data optimization is performed using a method most suitable for the sample, e.g., normalization with median values, normalization with z-scores, setting the maximum and the minimum values, or log transformation.

Means for Outputting List of Variables 605 displays a list of variables in the sample data to be subjected to cluster analysis.

Subsequently, using the function of Variable Selection Means 606, a user selects variables from the variables displayed by Means for Outputting List of Variables 605.

The selection of variables by Variable Selection Means 606 allows free selection of a single or a plurality of particular variables. Since, usually, there are a large number of candidates for variables, the Means 606 allows the user selection of any variables from them.

Once the user has selected particular variables, this information is input into Means for Generating Sample Data Files for Identification 607 together with sample data. Then, sample data files for identification is generated by this Generating Means 607.

Subsequently, the resultant data files of clusters are sent to Identification Means 608, which evaluates the degree of separation of clusters. The evaluation formula to evaluate the degree of separation of clusters may be defined in various manners.

The results of evaluation of the degree of separation made by the Identification Means 608 are passed to Means for Classifying Clusters 609. Then, the Classifying Means 609 inputs the identification results by the Identification Means 608 and decides most appropriate cluster classification referring to the identification conditions set in Means for Setting Identification Conditions 612. If conditions for continuation/termination of cluster classification have been set, the Classifying Means 609 judges the continuation or termination of cluster classification. If conditions for continuation/termination of cluster classification have not been set, the Classifying Means 609 allows the user to judge the continuation or termination of cluster classification. If the Classifying Means 609 has decided to continue cluster classification, it outputs the most appropriate cluster classification obtained from the processing of that time and a signal announcing that cluster classification is continued. This signal will work later as an instruction that processing of cluster classification must be returned to the Means for Outputting List of Variables 605 after the processing by Means for Editing Dendrograms 611.

On the other hand, if the Classifying Means 609 has decided to terminate cluster classification, it specifies the most appropriate cluster at that stage and outputs a signal

announcing that cluster classification is terminated. This signal will work later as an instruction to terminate the processing of cluster classification after the processing by Means for Editing Dendrograms 611.

After completion of the processing by the Classifying Means 609, the processing by Means for Generating Dendrograms 610 starts. The Generating Means 610 inputs the cluster classification decided by the Classifying Means 609, and displays a dendrogram based on the above cluster classification and characters of the variables pertaining to the cluster classification. When the Generating Means 610 has generated this cluster classification dendrogram, the user becomes able to grasp the current state of cluster classification visually. While generating the dendrogram, the Generating Means 610 displays the amounts of gene expression which are basis for the generation of the dendrogram so that the user can grasp the amounts visually. Then, Means for Editing Dendrograms 611 allows the user on the screen of a display device to edit (i.e., make additions, changes, or deletions to) the dendrogram generated by the Generating Means 610. The addition, change, or deletion of cluster classification is performed by the user with a processing instruction input device on the screen. For example, certain clusters may be designated, and variables of clusters to be classified further at a lower level may be indicated; or a plurality of clusters may be amalgamated. Alternatively, branches of a certain cluster classification may be deleted. While providing various tools to support the user's editing operation, Means for Editing Dendrograms 611 reads the meaning of the user's editing operation and automatically corrects data files of individual clusters accordingly. Furthermore, Means for Editing Dendrograms 611 preferably presents the judgment of Cluster Classifying Means 609 to continue or terminate the cluster classification and allows the user to input his/her final decision.

If it is decided that repeated processing of cluster classification should be continued, the processing is returned to Means for Outputting List of Variables 605. Then, the above-described processing performed by Means for Outputting List of Variables 605 through Means for Editing Dendrograms 611 is repeated.

From the thus analyzed data, whether the test samples are under ischemic conditions or not can be judged by checking into which cluster (i.e., ischemic cluster or normal cluster)

they have been classified.

5. Method of Screening for Ischemic Condition-Improving Drugs or Therapeutics for Ischemic Diseases

It is possible to screen for ischemic condition-improving drugs or therapeutics for ischemic diseases using as an indicator the expression levels of the genes which were revealed in the invention to show increased expression levels under ischemic conditions. Briefly, (a) whether or not the expression levels of the above genes return to their expression levels in a normal tissue, or (b) whether or not the expression profile of a gene group comprising a plurality of above genes returns to the normal expression profile in a normal tissue, by the administration of a drug to a test animal or test cell is examined. When the expression levels of the above genes have returned to their expression levels in a normal tissue as a result of the administration of the drug, the drug is evaluated as a candidate substance for an ischemic condition-improving drug or therapeutic for ischemic diseases.

BRIEF DESCRIPTION OF DRAWINGS

Fig. 1 shows procedures to identify genes of which expression levels change under ischemic conditions using a DNA chip.

Fig. 2 shows one embodiment of the DNA chip of the invention for examining ischemic conditions, and predicted results when test samples from patients have been hybridized on the DNA chip.

Fig. 3 shows one embodiment of synthetic-type DNA chip and procedures to detect genes in a test sample using the DNA chip.

Fig. 4 shows procedures to measure the expression levels of genes using GeneChip™.

Fig. 5 is a block diagram showing an ischemic condition identification system.

Fig. 6 is a flow chart showing an example of processing by an ischemic condition identification program.

Legend

501: CPU; 502: ROM; 503: RAM, 504: Input Unit; 505: Sending/Receiving Unit; 506: Output Unit; 507: HDD, 508: CD-ROM Drive; 509: CD-ROM; 510: Database; 601: Cluster Analysis Device; 602: Means for External Database Searching and Data Input; 603: Sample Data Storage Means; 604: Data Optimizing Means; 605: Means for Outputting List of Variables; 606: Variable Selection Means; 607: Means for Generating Sample Data File for Evaluation; 608: Evaluation Means; 609: Means for Classifying Clusters; 610: Means for Generating Dendrograms; 611: Means for Editing Dendrograms; 612: Means for Setting Evaluation Conditions.

BEST MODE FOR CARRYING OUT THE INVENTION

Hereinbelow, the present invention will be described more specifically with reference to the following Example, which should not be construed as limiting the scope of the invention.

Example 1. Identification of Genes Expressed under Ischemic Conditions using a DNA Chip

Genes expressed under ischemic conditions were identified by the procedures as shown in Fig. 4 using GeneChip System™ (Affymetrix). Briefly, 8-10 week-old bcl BLACK mice were anesthetized with inflane. Then, the carotid artery on both sides was exposed and ligated for 20 min to block the blood flow to thereby generate ischemic conditions in mice. Then, the blood flow was restored, and mice were slaughtered at 0, 2, 6, 12 and 24 hours after the blood flow restoration. The hippocampus on both sides was removed and immediately stored frozen. Subsequently, approximately 2 μ g of poly(A)+mRNA was extracted from 1 g of a frozen sample. Then, cDNA was synthesized with a reverse transcriptase. The resultant cDNA was transcribed *in vitro* to thereby obtain biotin-labeled cRNA, which was treated with heat in the presence of Mg^{2+} for fragmentation into about 50-mer fragments. Internal reference was labeled and added to the sample. Then, the sample was poured into GeneChip™ Mu6500 (Affymetrix). After hybridization in an oven, the chip was washed in Fluidic station. Then, information of the chip was read by a GeneArray scanner. The data obtained was processed and analyzed using a

bioinformatics system. The results are shown in Tables 1 through 3. Table 1 shows genes of which transcription levels increased n-fold where n is 2 or more but less than 5 within 24 hours when the transcription level at 0 hour is taken as 1. Table 2 shows genes of which transcription levels increased n-fold where n is 5 or more but less than 10. Table 3 shows genes of which transcription levels increased more than 10-fold. The accession numbers refers to GenBank accession numbers.

Table 1.

Genes of Which Transcription Levels Increased n-Fold ($2 \leqq n < 5$)

ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)	ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)
M77167	SEQ ID NO.1	SEQ ID NO.2	AA103541	SEQ ID NO.465	–
AA068780	SEQ ID NO.3	–	U41736	SEQ ID NO.466	SEQ ID NO.467
M33212	SEQ ID NO.4	SEQ ID NO.5	AA103768	SEQ ID NO.468	–
D76440	SEQ ID NO.6	SEQ ID NO.7	U43900	SEQ ID NO.469	SEQ ID NO.470
M35131	SEQ ID NO.8	SEQ ID NO.9	AA140170	SEQ ID NO.471	–
AA050852	SEQ ID NO.10	–	AA104086	SEQ ID NO.472	–
W47946	SEQ ID NO.11	–	U43512	SEQ ID NO.473	–
W48104	SEQ ID NO.12	–	AA104459	SEQ ID NO.474	–
U09928	SEQ ID NO.13	–	U39545	SEQ ID NO.475	SEQ ID NO.476
M29464	SEQ ID NO.14	SEQ ID NO.15	U64445	SEQ ID NO.477	SEQ ID NO.478
X61434	SEQ ID NO.16	–	D86949	SEQ ID NO.479	SEQ ID NO.480
M90316	SEQ ID NO.17	SEQ ID NO.18	Z70023	SEQ ID NO.481	SEQ ID NO.482
M94632	SEQ ID NO.19	SEQ ID NO.20	U71202	SEQ ID NO.483	SEQ ID NO.484
W41501	SEQ ID NO.21	–	X81059	SEQ ID NO.485	SEQ ID NO.486
M20632	SEQ ID NO.22	SEQ ID NO.23	U73200	SEQ ID NO.487	SEQ ID NO.488
M13521	SEQ ID NO.24	–	U69270	SEQ ID NO.489	SEQ ID NO.490
M21285	SEQ ID NO.25	–	AA106077	SEQ ID NO.491	–
M96163	SEQ ID NO.26	–	AA106116	SEQ ID NO.492	–
X04017	SEQ ID NO.27	SEQ ID NO.28	U31758	SEQ ID NO.493	SEQ ID NO.494
M64086	SEQ ID NO.29	SEQ ID NO.30	AA106224	SEQ ID NO.495	–
W54584	SEQ ID NO.31	–	X80502	SEQ ID NO.496	SEQ ID NO.497
L12447	SEQ ID NO.32	SEQ ID NO.33	U76832	SEQ ID NO.498	SEQ ID NO.499
X62622	SEQ ID NO.34	SEQ ID NO.35	U30840	SEQ ID NO.500	SEQ ID NO.501
Z30970	SEQ ID NO.36	SEQ ID NO.37	AA107388	SEQ ID NO.502	–
M23568	SEQ ID NO.38	SEQ ID NO.39	U26459	SEQ ID NO.503	SEQ ID NO.504
J03520	SEQ ID NO.40	SEQ ID NO.41	U29055	SEQ ID NO.505	SEQ ID NO.506
M94087	SEQ ID NO.42	SEQ ID NO.43	U39302	SEQ ID NO.507	SEQ ID NO.508
L37092	SEQ ID NO.44	SEQ ID NO.45	U77630	SEQ ID NO.509	SEQ ID NO.510
M22479	SEQ ID NO.46	SEQ ID NO.47	U39473	SEQ ID NO.511	SEQ ID NO.512

AA041634	SEQ ID NO.48	–	X91043	SEQ ID NO.513	SEQ ID NO.514
M95200	SEQ ID NO.49	SEQ ID NO.50	D86603	SEQ ID NO.515	SEQ ID NO.516
M62867	SEQ ID NO.51	SEQ ID NO.52	D88792	SEQ ID NO.517	SEQ ID NO.518
AA137883	SEQ ID NO.53	–	D88793	SEQ ID NO.519	SEQ ID NO.520
L36611	SEQ ID NO.54	–	D38218	SEQ ID NO.521	SEQ ID NO.522
L19311	SEQ ID NO.55	SEQ ID NO.56	U60330	SEQ ID NO.523	SEQ ID NO.524
X74438	SEQ ID NO.57	SEQ ID NO.58	AA108925	SEQ ID NO.525	–
AA034714	SEQ ID NO.59	–	U66865	SEQ ID NO.526	SEQ ID NO.527
D32167	SEQ ID NO.60	SEQ ID NO.61	Y07836	SEQ ID NO.528	SEQ ID NO.529
U24700	SEQ ID NO.62	SEQ ID NO.63	AA110543	SEQ ID NO.530	–
U07602	SEQ ID NO.64	–	AA117787	SEQ ID NO.531	–
U66620	SEQ ID NO.65	SEQ ID NO.66	D30785	SEQ ID NO.532	SEQ ID NO.533
L25126	SEQ ID NO.67	SEQ ID NO.68	U69488	SEQ ID NO.534	–
U02982	SEQ ID NO.69	SEQ ID NO.70	Z11997	SEQ ID NO.535	SEQ ID NO.536
AA155371	SEQ ID NO.71	–	D38613	SEQ ID NO.537	SEQ ID NO.538
U16740	SEQ ID NO.72	–	X92352	SEQ ID NO.539	SEQ ID NO.540
U16741	SEQ ID NO.73	SEQ ID NO.74	AA111168	SEQ ID NO.541	–
W58941	SEQ ID NO.75	–	AA111212	SEQ ID NO.542	–
U22948	SEQ ID NO.76	SEQ ID NO.77	W07963	SEQ ID NO.543	–
U23789	SEQ ID NO.78	SEQ ID NO.79	AA114781	SEQ ID NO.544	–
X68951	SEQ ID NO.80	SEQ ID NO.81	W08033	SEQ ID NO.545	–
AA144400	SEQ ID NO.82	–	AA117064	SEQ ID NO.546	–
U06670	SEQ ID NO.83	SEQ ID NO.84	AA117128	SEQ ID NO.547	–
U09399	SEQ ID NO.85	SEQ ID NO.86	AA162093	SEQ ID NO.548	–
U10118	SEQ ID NO.87	–	U10871	SEQ ID NO.549	SEQ ID NO.550
L17069	SEQ ID NO.88	–	AA120173	SEQ ID NO.551	–
W62091	SEQ ID NO.89	–	AA123463	SEQ ID NO.552	–
U09419	SEQ ID NO.90	SEQ ID NO.91	W07927	SEQ ID NO.553	–
AA041651	SEQ ID NO.92	–	AA118546	SEQ ID NO.554	–
AA051063	SEQ ID NO.93	–	AA118294	SEQ ID NO.555	–
W62420	SEQ ID NO.94	–	AA169951	SEQ ID NO.556	–
AA110896	SEQ ID NO.95	–	AA146194	SEQ ID NO.557	–
M60057	SEQ ID NO.96	–	W08228	SEQ ID NO.558	–
U33840	SEQ ID NO.97	SEQ ID NO.98	W33440	SEQ ID NO.559	–
W63835	SEQ ID NO.99	–	AA119959	SEQ ID NO.560	–
W63876	SEQ ID NO.100	–	W08326	SEQ ID NO.561	–
AA050268	SEQ ID NO.101	SEQ ID NO.	AA125425	SEQ ID NO.562	–
W63974	SEQ ID NO.102	SEQ ID NO.	AA137580	SEQ ID NO.563	–
D14571	SEQ ID NO.103	SEQ ID NO.104	AA138848	SEQ ID NO.564	–
L20450	SEQ ID NO.105	SEQ ID NO.106	AA145188	SEQ ID NO.565	–
W64413	SEQ ID NO.107	–	AA145160	SEQ ID NO.566	–
D16262	SEQ ID NO.108	SEQ ID NO.109	W13461	SEQ ID NO.567	–
AA120109	SEQ ID NO.110	–	W13878	SEQ ID NO.568	–
X61432	SEQ ID NO.111	SEQ ID NO.112	U35623	SEQ ID NO.569	SEQ ID NO.570

D10011	SEQ ID NO.113	SEQ ID NO.114	W77121	SEQ ID NO.571	–
L16953	SEQ ID NO.115	SEQ ID NO.116	W14370	SEQ ID NO.572	–
D29016	SEQ ID NO.117	SEQ ID NO.118	AA152884	SEQ ID NO.573	–
W65920	SEQ ID NO.119	–	AA023591	SEQ ID NO.574	–
L39123	SEQ ID NO.120	SEQ ID NO.121	AA155191	SEQ ID NO.575	–
U16162	SEQ ID NO.122	–	AA050733	SEQ ID NO.576	–
AA002504	SEQ ID NO.123	–	AA096645	SEQ ID NO.577	–
U20326	SEQ ID NO.124	SEQ ID NO.125	AA165847	SEQ ID NO.578	–
U27830	SEQ ID NO.126	SEQ ID NO.127	AA170223	SEQ ID NO.579	–
U24703	SEQ ID NO.128	SEQ ID NO.129	AA170375	SEQ ID NO.580	–
W70782	SEQ ID NO.130	–	Z25524	SEQ ID NO.581	SEQ ID NO.582
U28217	SEQ ID NO.131	SEQ ID NO.132	X16857	SEQ ID NO.583	–
AA105294	SEQ ID NO.133	–	D90151	SEQ ID NO.584	SEQ ID NO.585
U28138	SEQ ID NO.134	–	D00925	SEQ ID NO.586	SEQ ID NO.587
U29396	SEQ ID NO.135	SEQ ID NO.136	ditto	ditto	SEQ ID NO.588
L29441	SEQ ID NO.137	SEQ ID NO.138	L76223	SEQ ID NO.589	–
D10061	SEQ ID NO.139	SEQ ID NO.140	D00613	SEQ ID NO.590	SEQ ID NO.591
W75403	SEQ ID NO.141	–	W13502	SEQ ID NO.592	–
M73696	SEQ ID NO.142	SEQ ID NO.143	D90344	SEQ ID NO.593	SEQ ID NO.594
W75531	SEQ ID NO.144	–	AA138226	SEQ ID NO.595	–
W75616	SEQ ID NO.145	–	AA003230	SEQ ID NO.596	–
L05439	SEQ ID NO.146	–	W30651	SEQ ID NO.597	–
D16432	SEQ ID NO.147	SEQ ID NO.148	M22531	SEQ ID NO.598	SEQ ID NO.599
L21027	SEQ ID NO.149	SEQ ID NO.150	AA118758	SEQ ID NO.600	–
AA009150	SEQ ID NO.151	–	AA027739	SEQ ID NO.601	–
U13262	SEQ ID NO.152	SEQ ID NO.153	AA168362	SEQ ID NO.602	–
W78604	SEQ ID NO.154	–	D78645	SEQ ID NO.603	SEQ ID NO.604
X56045	SEQ ID NO.155	SEQ ID NO.156	M31131	SEQ ID NO.605	SEQ ID NO.606
U36277	SEQ ID NO.157	SEQ ID NO.158	AA153320	SEQ ID NO.607	–
W82115	SEQ ID NO.159	–	W34066	SEQ ID NO.608	–
M69293	SEQ ID NO.160	–	W10606	SEQ ID NO.609	–
U36760	SEQ ID NO.161	SEQ ID NO.162	W20828	SEQ ID NO.610	–
U35142	SEQ ID NO.163	SEQ ID NO.164	L08115	SEQ ID NO.611	SEQ ID NO.612
U14420	SEQ ID NO.165	SEQ ID NO.166	Z31554	SEQ ID NO.613	SEQ ID NO.614
Y00769	SEQ ID NO.167	SEQ ID NO.168	W41861	SEQ ID NO.615	–
U37720	SEQ ID NO.169	SEQ ID NO.170	W71125	SEQ ID NO.616	–
U35368	SEQ ID NO.171	SEQ ID NO.172	AA000961	SEQ ID NO.617	–
W89939	SEQ ID NO.173	–	W82998	SEQ ID NO.618	–
W37000	SEQ ID NO.174	–	W11954	SEQ ID NO.619	–
W97373	SEQ ID NO.175	–	U43844	SEQ ID NO.620	–
W96831	SEQ ID NO.176	–	AA008650	SEQ ID NO.621	–
W82209	SEQ ID NO.177	–	W18601	SEQ ID NO.622	–
AA009169	SEQ ID NO.178	–	W30116	SEQ ID NO.623	–
U31967	SEQ ID NO.179	SEQ ID NO.180	U53456	SEQ ID NO.624	SEQ ID NO.625

D28117	SEQ ID NO.181	SEQ ID NO.182	M27073	SEQ ID NO.626	SEQ ID NO.627
U34691	SEQ ID NO.183	SEQ ID NO.184	W12720	SEQ ID NO.628	—
U32329	SEQ ID NO.185	SEQ ID NO.186	M76131	SEQ ID NO.629	SEQ ID NO.630
U10903	SEQ ID NO.187	SEQ ID NO.188	U39192	SEQ ID NO.631	—
D49949	SEQ ID NO.189	SEQ ID NO.190	W54288	SEQ ID NO.632	—
D49429	SEQ ID NO.191	SEQ ID NO.192	M76124	SEQ ID NO.633	SEQ ID NO.634
U48397	SEQ ID NO.193	SEQ ID NO.194	W12946	SEQ ID NO.635	—
D00622	SEQ ID NO.195	—	AA024288	SEQ ID NO.636	—
U38940	SEQ ID NO.196	SEQ ID NO.197	W13412	SEQ ID NO.637	—
X75313	SEQ ID NO.198	SEQ ID NO.199	AA002852	SEQ ID NO.638	—
AA008683	SEQ ID NO.200	—	W15789	SEQ ID NO.639	—
U37799	SEQ ID NO.201	SEQ ID NO.202	L39879	SEQ ID NO.640	—
U43918	SEQ ID NO.203	SEQ ID NO.204	M61215	SEQ ID NO.641	SEQ ID NO.642
AA039108	SEQ ID NO.205	—	M70641	SEQ ID NO.643	—
L10244	SEQ ID NO.206	SEQ ID NO.207	D86729	SEQ ID NO.644	SEQ ID NO.645
D78647	SEQ ID NO.208	SEQ ID NO.209	L01640	SEQ ID NO.646	SEQ ID NO.647
U36220	SEQ ID NO.210	SEQ ID NO.211	M92378	SEQ ID NO.648	—
AA051121	SEQ ID NO.212	—	K01347	SEQ ID NO.649	SEQ ID NO.650
AA003458	SEQ ID NO.213	—	M23384	SEQ ID NO.651	SEQ ID NO.652
AA003554	SEQ ID NO.214	—	W77613	SEQ ID NO.653	—
U46934	SEQ ID NO.215	SEQ ID NO.216	W18385	SEQ ID NO.654	—
D50086	SEQ ID NO.217	SEQ ID NO.218	W53188	SEQ ID NO.655	—
AA007899	SEQ ID NO.219	—	J05277	SEQ ID NO.656	SEQ ID NO.657
AA008133	SEQ ID NO.220	—	W30499	SEQ ID NO.658	—
X58876	SEQ ID NO.221	SEQ ID NO.222	W67014	SEQ ID NO.659	—
D28530	SEQ ID NO.223	SEQ ID NO.224	M31885	SEQ ID NO.660	—
AA048974	SEQ ID NO.225	—	W54228	SEQ ID NO.661	—
AA120290	SEQ ID NO.226	—	W29756	SEQ ID NO.662	—
L40406	SEQ ID NO.227	SEQ ID NO.228	W30609	SEQ ID NO.663	—
U27340	SEQ ID NO.229	SEQ ID NO.230	D17666	SEQ ID NO.664	—
AA013581	SEQ ID NO.231	—	AA033408	SEQ ID NO.665	—
AA013830	SEQ ID NO.232	—	W33728	SEQ ID NO.666	—
Z37110	SEQ ID NO.233	SEQ ID NO.234	W74850	SEQ ID NO.667	—
AA119078	SEQ ID NO.235	—	AA119191	SEQ ID NO.668	—
Z67746	SEQ ID NO.236	SEQ ID NO.237	J05205	SEQ ID NO.669	SEQ ID NO.670
U51196	SEQ ID NO.238	SEQ ID NO.239	X06746	SEQ ID NO.671	SEQ ID NO.672
AA118878	SEQ ID NO.240	—	J02870	SEQ ID NO.673	SEQ ID NO.674
X89650	SEQ ID NO.241	SEQ ID NO.242	AA124273	SEQ ID NO.675	—
U07617	SEQ ID NO.243	SEQ ID NO.244	AA008245	SEQ ID NO.676	—
U37351	SEQ ID NO.245	SEQ ID NO.246	AA106492	SEQ ID NO.677	—
Y08640	SEQ ID NO.247	SEQ ID NO.248	X07997	SEQ ID NO.678	—
U48797	SEQ ID NO.249	SEQ ID NO.250	M21041	SEQ ID NO.679	SEQ ID NO.680
AA020620	SEQ ID NO.251	—	W36757	SEQ ID NO.681	—
X01023	SEQ ID NO.252	SEQ ID NO.253	V00835	SEQ ID NO.682	—

U34920	SEQ ID NO.254	SEQ ID NO.255	Y00305	SEQ ID NO.683	SEQ ID NO.684
U17698	SEQ ID NO.256	SEQ ID NO.257	M64640	SEQ ID NO.685	–
X54352	SEQ ID NO.258	SEQ ID NO.259	W42216	SEQ ID NO.686	–
U25708	SEQ ID NO.260	SEQ ID NO.261	U73478	SEQ ID NO.687	SEQ ID NO.688
X15267	SEQ ID NO.262	SEQ ID NO.263	J00398	SEQ ID NO.689	–
X66091	SEQ ID NO.264	–	M68899	SEQ ID NO.690	SEQ ID NO.691
X55573	SEQ ID NO.265	SEQ ID NO.266	W34232	SEQ ID NO.692	–
AA119194	SEQ ID NO.267	–	D38077	SEQ ID NO.693	–
X13586	SEQ ID NO.268	SEQ ID NO.269	W29533	SEQ ID NO.694	–
Z31553	SEQ ID NO.270	SEQ ID NO.271	W29462	SEQ ID NO.695	–
Z31399	SEQ ID NO.272	SEQ ID NO.273	W62036	SEQ ID NO.696	–
AA108956	SEQ ID NO.274	–	W77226	SEQ ID NO.697	–
Z31557	SEQ ID NO.275	SEQ ID NO.276	D14340	SEQ ID NO.698	SEQ ID NO.699
AA032948	SEQ ID NO.277	–	M83118	SEQ ID NO.700	SEQ ID NO.701
Z12302	SEQ ID NO.278	SEQ ID NO.279	M21952	SEQ ID NO.702	SEQ ID NO.703
X60367	SEQ ID NO.280	SEQ ID NO.281	W11666	SEQ ID NO.704	–
X64713	SEQ ID NO.282	SEQ ID NO.283	AA015415	SEQ ID NO.705	–
AA035938	SEQ ID NO.284	–	AA097018	SEQ ID NO.706	–
AA036445	SEQ ID NO.285	–	M27796	SEQ ID NO.707	SEQ ID NO.708
X17502	SEQ ID NO.286	SEQ ID NO.287	AA016424	SEQ ID NO.709	–
X79233	SEQ ID NO.288	SEQ ID NO.289	W29543	SEQ ID NO.710	–
X13135	SEQ ID NO.290	–	AA028728	SEQ ID NO.711	–
Z22593	SEQ ID NO.291	SEQ ID NO.292	K00988	SEQ ID NO.712	–
X90875	SEQ ID NO.293	SEQ ID NO.294	AA170547	SEQ ID NO.713	–
Z36270	SEQ ID NO.295	SEQ ID NO.296	W20613	SEQ ID NO.714	–
X67056	SEQ ID NO.297	SEQ ID NO.298	AA163687	SEQ ID NO.715	–
X13605	SEQ ID NO.299	SEQ ID NO.300	M93422	SEQ ID NO.716	SEQ ID NO.717
W89699	SEQ ID NO.301	–	W12548	SEQ ID NO.718	–
X62669	SEQ ID NO.302	–	AA145181	SEQ ID NO.719	–
AA049790	SEQ ID NO.303	–	AA144057	SEQ ID NO.720	–
M33227	SEQ ID NO.304	SEQ ID NO.305	AA138791	SEQ ID NO.721	–
AA060064	SEQ ID NO.306	–	L27453	SEQ ID NO.722	SEQ ID NO.723
X70393	SEQ ID NO.307	SEQ ID NO.308	AA119571	SEQ ID NO.724	–
AA050789	SEQ ID NO.309	–	AA117227	SEQ ID NO.725	–
AA051486	SEQ ID NO.310	–	U03645	SEQ ID NO.726	SEQ ID NO.727
X61147	SEQ ID NO.311	SEQ ID NO.312	Y07711	SEQ ID NO.728	SEQ ID NO.729
AA058163	SEQ ID NO.313	–	D17571	SEQ ID NO.730	SEQ ID NO.731
W14434	SEQ ID NO.314	–	U68058	SEQ ID NO.732	SEQ ID NO.733
W82720	SEQ ID NO.315	–	U66202	SEQ ID NO.734	SEQ ID NO.735
W89958	SEQ ID NO.316	–	U56651	SEQ ID NO.736	–
AA107471	SEQ ID NO.317	–	AA110711	SEQ ID NO.737	–
AA028547	SEQ ID NO.318	–	X92122	SEQ ID NO.738	SEQ ID NO.739
X60831	SEQ ID NO.319	SEQ ID NO.320	X87257	SEQ ID NO.740	–
Z49085	SEQ ID NO.321	SEQ ID NO.322	U69695	SEQ ID NO.741	SEQ ID NO.742

X05640	SEQ ID NO.323	–	X64550	SEQ ID NO.743	–
AA059967	SEQ ID NO.324	–	U59463	SEQ ID NO.744	SEQ ID NO.745
AA163305	SEQ ID NO.325	–	AA096532	SEQ ID NO.746	–
X64837	SEQ ID NO.326	SEQ ID NO.327	W85446	SEQ ID NO.747	–
AA060704	SEQ ID NO.328	–	X94444	SEQ ID NO.748	SEQ ID NO.749
X17320	SEQ ID NO.329	SEQ ID NO.330	AA080097	SEQ ID NO.750	–
X65553	SEQ ID NO.331	SEQ ID NO.332	D32132	SEQ ID NO.751	–
X51468	SEQ ID NO.333	–	Z71268	SEQ ID NO.752	SEQ ID NO.753
X14425	SEQ ID NO.334	SEQ ID NO.335	U28419	SEQ ID NO.754	–
AA061707	SEQ ID NO.336	–	X99641	SEQ ID NO.755	SEQ ID NO.756
X95403	SEQ ID NO.337	SEQ ID NO.338	AA072611	SEQ ID NO.757	–
AA139612	SEQ ID NO.339	–	X90647	SEQ ID NO.758	–
X57277	SEQ ID NO.340	SEQ ID NO.341	X95580	SEQ ID NO.759	SEQ ID NO.760
X70067	SEQ ID NO.342	SEQ ID NO.343	U46187	SEQ ID NO.761	SEQ ID NO.762
U05245	SEQ ID NO.344	–	U43085	SEQ ID NO.763	SEQ ID NO.764
AA062237	SEQ ID NO.345	–	L29479	SEQ ID NO.765	SEQ ID NO.766
AA063841	SEQ ID NO.346	–	U20892	SEQ ID NO.767	–
X89222	SEQ ID NO.347	SEQ ID NO.348	AA064467	SEQ ID NO.768	–
AA064004	SEQ ID NO.349	–	W09407	SEQ ID NO.769	–
AA087787	SEQ ID NO.350	–	AA015354	SEQ ID NO.770	–
AA064355	SEQ ID NO.351	–	Z31370	SEQ ID NO.771	SEQ ID NO.772
X68837	SEQ ID NO.352	SEQ ID NO.353	X57349	SEQ ID NO.773	SEQ ID NO.774
Z19579	SEQ ID NO.354	SEQ ID NO.355	X81464	SEQ ID NO.775	SEQ ID NO.776
Z19581	SEQ ID NO.356	SEQ ID NO.357	W89293	SEQ ID NO.777	–
X65657	SEQ ID NO.358	SEQ ID NO.359	X58523	SEQ ID NO.778	SEQ ID NO.779
X16319	SEQ ID NO.360	SEQ ID NO.361	X53476	SEQ ID NO.780	SEQ ID NO.781
X78304	SEQ ID NO.362	SEQ ID NO.363	AA044535	SEQ ID NO.782	–
X81987	SEQ ID NO.364	SEQ ID NO.365	AA037964	SEQ ID NO.783	–
U52197	SEQ ID NO.366	SEQ ID NO.367	X81632	SEQ ID NO.784	–
AA065652	SEQ ID NO.368	–	AA035984	SEQ ID NO.785	–
U23921	SEQ ID NO.369	SEQ ID NO.370	AA035915	SEQ ID NO.786	–
AA066333	SEQ ID NO.371	–	X61800	SEQ ID NO.787	SEQ ID NO.788
U38501	SEQ ID NO.372	–	U03457	SEQ ID NO.789	SEQ ID NO.790
W46019	SEQ ID NO.373	–	AA023390	SEQ ID NO.791	–
D49744	SEQ ID NO.374	SEQ ID NO.375	AA064208	SEQ ID NO.792	–
U58497	SEQ ID NO.376	SEQ ID NO.377	L75822	SEQ ID NO.793	SEQ ID NO.794
U38252	SEQ ID NO.378	SEQ ID NO.379	U49739	SEQ ID NO.795	SEQ ID NO.796
U40930	SEQ ID NO.380	SEQ ID NO.381	U41636	SEQ ID NO.797	SEQ ID NO.798
AA118729	SEQ ID NO.382	–	X70296	SEQ ID NO.799	SEQ ID NO.800
AA068057	SEQ ID NO.383	–	AA161905	SEQ ID NO.801	–
U47104	SEQ ID NO.384	SEQ ID NO.385	AA016858	SEQ ID NO.802	–
M97635	SEQ ID NO.386	–	U50734	SEQ ID NO.803	SEQ ID NO.804
U58883	SEQ ID NO.387	SEQ ID NO.388	U36993	SEQ ID NO.805	SEQ ID NO.806
D49654	SEQ ID NO.389	SEQ ID NO.390	AA013513	SEQ ID NO.807	–

D78641	SEQ ID NO.391	SEQ ID NO.392	U31992	SEQ ID NO.808	SEQ ID NO.809
U24428	SEQ ID NO.393	SEQ ID NO.394	AA007871	SEQ ID NO.810	–
L43326	SEQ ID NO.395	SEQ ID NO.396	U43317	SEQ ID NO.811	SEQ ID NO.812
U08378	SEQ ID NO.397	SEQ ID NO.398	U41465	SEQ ID NO.813	SEQ ID NO.814
D78382	SEQ ID NO.399	SEQ ID NO.400	AA000341	SEQ ID NO.815	–
AA073296	SEQ ID NO.401	–	U39738	SEQ ID NO.816	SEQ ID NO.817
D90173	SEQ ID NO.402	SEQ ID NO.403	U28728	SEQ ID NO.818	SEQ ID NO.819
U62483	SEQ ID NO.404	SEQ ID NO.405	M85151	SEQ ID NO.820	SEQ ID NO.821
D50523	SEQ ID NO.406	–	U04827	SEQ ID NO.822	–
U63323	SEQ ID NO.407	SEQ ID NO.408	U35885	SEQ ID NO.823	SEQ ID NO.824
AA079926	SEQ ID NO.409	–	X77557	SEQ ID NO.825	SEQ ID NO.826
X99963	SEQ ID NO.410	SEQ ID NO.411	D14883	SEQ ID NO.827	SEQ ID NO.828
X97281	SEQ ID NO.412	SEQ ID NO.413	U28068	SEQ ID NO.829	–
X92411	SEQ ID NO.414	SEQ ID NO.415	W65511	SEQ ID NO.830	–
AA087332	SEQ ID NO.416	–	U17961	SEQ ID NO.831	SEQ ID NO.832
X80232	SEQ ID NO.417	SEQ ID NO.418	L32973	SEQ ID NO.833	SEQ ID NO.834
U49350	SEQ ID NO.419	SEQ ID NO.420	D16333	SEQ ID NO.835	SEQ ID NO.836
U49351	SEQ ID NO.421	SEQ ID NO.422	Z67747	SEQ ID NO.837	–
AA097087	SEQ ID NO.423	–	D16580	SEQ ID NO.838	SEQ ID NO.839
AA087616	SEQ ID NO.424	–	M18776	SEQ ID NO.840	SEQ ID NO.841
D73368	SEQ ID NO.425	SEQ ID NO.426	M91458	SEQ ID NO.842	SEQ ID NO.843
U48804	SEQ ID NO.427	SEQ ID NO.428	AA120695	SEQ ID NO.844	–
AA087986	SEQ ID NO.429	–	L29503	SEQ ID NO.845	–
AA088003	SEQ ID NO.430	–	M25811	SEQ ID NO.846	SEQ ID NO.847
Y07685	SEQ ID NO.431	SEQ ID NO.432	M34141	SEQ ID NO.848	SEQ ID NO.849
Y07693	SEQ ID NO.433	SEQ ID NO.434	W45807	SEQ ID NO.850	–
Y07688	SEQ ID NO.435	SEQ ID NO.436	M63903	SEQ ID NO.851	SEQ ID NO.852
D67015	SEQ ID NO.437	SEQ ID NO.438	U10119	SEQ ID NO.853	SEQ ID NO.854
U67874	SEQ ID NO.439	SEQ ID NO.440	L35556	SEQ ID NO.855	SEQ ID NO.856
D86344	SEQ ID NO.441	SEQ ID NO.442	X98096	SEQ ID NO.857	SEQ ID NO.858
AA097203	SEQ ID NO.443	–	AA166601	SEQ ID NO.1038	–
Y08135	SEQ ID NO.444	SEQ ID NO.445	AA125097	SEQ ID NO.1039	–
U42384	SEQ ID NO.446	SEQ ID NO.447	AA168363	SEQ ID NO.1040	–
D87899	SEQ ID NO.448	SEQ ID NO.449	U19582	SEQ ID NO.1041	SEQ ID NO.1042
D87901	SEQ ID NO.450	SEQ ID NO.451	L22482	SEQ ID NO.1043	SEQ ID NO.1044
D87902	SEQ ID NO.452	SEQ ID NO.453	W63876	SEQ ID NO.1045	–
D87903	SEQ ID NO.454	SEQ ID NO.455	U11274	SEQ ID NO.1046	SEQ ID NO.1047
X96859	SEQ ID NO.456	–	W66636	SEQ ID NO.1048	–
U56909	SEQ ID NO.457	SEQ ID NO.458	D42124	SEQ ID NO.1049	SEQ ID NO.1050
AA098588	SEQ ID NO.459	–	U41765	SEQ ID NO.1051	SEQ ID NO.1052
X70298	SEQ ID NO.460	SEQ ID NO.461	X02452	SEQ ID NO.1053	–
U62021	SEQ ID NO.462	–	X76653	SEQ ID NO.1054	SEQ ID NO.1055
D90225	SEQ ID NO.463	SEQ ID NO.464	W13425	SEQ ID NO.1056	–

Table 2.

Genes of Which Transcription Levels Increased n-Fold ($5 \leqq n < 10$)

ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)	ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)
X16995	SEQ ID NO.859	–	M90364	SEQ ID NO.917	SEQ ID NO.918
W78418	SEQ ID NO.860	–	M16465	SEQ ID NO.919	SEQ ID NO.920
L10106	SEQ ID NO.861	SEQ ID NO.862	W13425	SEQ ID NO.921	–
AA060409	SEQ ID NO.863	–	M62766	SEQ ID NO.922	SEQ ID NO.923
M57999	SEQ ID NO.864	SEQ ID NO.865	U73744	SEQ ID NO.924	–
M26251	SEQ ID NO.866	SEQ ID NO.867	W29669	SEQ ID NO.925	–
M22432	SEQ ID NO.868	SEQ ID NO.869	W33838	SEQ ID NO.926	–
U00689	SEQ ID NO.870	SEQ ID NO.871	L21707	SEQ ID NO.927	SEQ ID NO.928
L22482	SEQ ID NO.872	SEQ ID NO.873	W34687	SEQ ID NO.929	–
U70662	SEQ ID NO.874	SEQ ID NO.875	AA105763	SEQ ID NO.930	–
AA064330	SEQ ID NO.876	–	M18194	SEQ ID NO.931	–
D55720	SEQ ID NO.877	SEQ ID NO.878	J04694	SEQ ID NO.932	SEQ ID NO.933
U37465	SEQ ID NO.879	SEQ ID NO.880	AA140026	SEQ ID NO.934	–
AA003413	SEQ ID NO.881	–	L35303	SEQ ID NO.935	SEQ ID NO.936
U43319	SEQ ID NO.882	SEQ ID NO.883	AA073986	SEQ ID NO.937	–
Z72000	SEQ ID NO.884	SEQ ID NO.885	X96618	SEQ ID NO.938	SEQ ID NO.939
X96639	SEQ ID NO.886	SEQ ID NO.887	X85983	SEQ ID NO.940	SEQ ID NO.941
X61940	SEQ ID NO.888	SEQ ID NO.889	X04367	SEQ ID NO.942	SEQ ID NO.943
X56518	SEQ ID NO.890	–	X07967	SEQ ID NO.944	SEQ ID NO.945
X94998	SEQ ID NO.891	SEQ ID NO.892	D90146	SEQ ID NO.946	–
X71642	SEQ ID NO.893	SEQ ID NO.894	AA014024	SEQ ID NO.947	–
X53584	SEQ ID NO.895	–	M58566	SEQ ID NO.948	SEQ ID NO.949
W89863	SEQ ID NO.896	–	AA044561	SEQ ID NO.950	–
X53257	SEQ ID NO.897	SEQ ID NO.898	W65899	SEQ ID NO.951	–
X76858	SEQ ID NO.899	SEQ ID NO.900	D10217	SEQ ID NO.952	SEQ ID NO.953
AA062131	SEQ ID NO.901	–	M55154	SEQ ID NO.954	SEQ ID NO.955
U12473	SEQ ID NO.902	SEQ ID NO.903	W46015	SEQ ID NO.956	–
M68859	SEQ ID NO.904	–	M63554	SEQ ID NO.957	SEQ ID NO.958
AA067929	SEQ ID NO.905	–	M13990	SEQ ID NO.959	–
U39066	SEQ ID NO.906	SEQ ID NO.907	W46084	SEQ ID NO.1057	–
U44088	SEQ ID NO.908	SEQ ID NO.909	AA048304	SEQ ID NO.1058	–
D87691	SEQ ID NO.910	SEQ ID NO.911	AA036574	SEQ ID NO.1059	–
U24160	SEQ ID NO.912	SEQ ID NO.913	AA030865	SEQ ID NO.1060	–
AA142376	SEQ ID NO.914	–	X51829	SEQ ID NO.1061	SEQ ID NO.1062
AA140150	SEQ ID NO.915	–	X63440	SEQ ID NO.1063	SEQ ID NO.1064
W29163	SEQ ID NO.916	–			

Table 3.

Genes of Which Transcription Levels Increased More Than 10-Fold

ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)	ACCESSION NO.	SEQ ID NO. (Nucleotide Sequence)	SEQ ID NO. (Amino Acid Sequence)
W49108	SEQ ID NO.960	–	J04115	SEQ ID NO.1000	SEQ ID NO.1001
M64292	SEQ ID NO.961	–	M22326	SEQ ID NO.1002	SEQ ID NO.1003
L25125	SEQ ID NO.962	SEQ ID NO.963	X14897	SEQ ID NO.1004	SEQ ID NO.1005
D49382	SEQ ID NO.964	SEQ ID NO.965	M88242	SEQ ID NO.1006	SEQ ID NO.1007
AA153519	SEQ ID NO.966	–	AA138777	SEQ ID NO.1008	–
AA013648	SEQ ID NO.967	–	AA061624	SEQ ID NO.1009	–
AA003990	SEQ ID NO.968	–	J00424	SEQ ID NO.1010	–
L20294	SEQ ID NO.969	SEQ ID NO.970	AA003162	SEQ ID NO.1011	–
W89900	SEQ ID NO.971	–	AA104477	SEQ ID NO.1012	–
U28656	SEQ ID NO.972	SEQ ID NO.973	J03236	SEQ ID NO.1013	SEQ ID NO.1014
U41805	SEQ ID NO.974	–	W33559	SEQ ID NO.1015	–
V00727	SEQ ID NO.975	–	AA041826	SEQ ID NO.1016	–
D78644	SEQ ID NO.976	–	L23971	SEQ ID NO.1017	SEQ ID NO.1018
X67083	SEQ ID NO.977	SEQ ID NO.978	AA117973	SEQ ID NO.1019	–
Z11911	SEQ ID NO.979	SEQ ID NO.980	AA117286	SEQ ID NO.1020	–
W17589	SEQ ID NO.981	–	U51037	SEQ ID NO.1021	SEQ ID NO.1022
X68273	SEQ ID NO.982	SEQ ID NO.983	U66203	SEQ ID NO.1023	SEQ ID NO.1024
X63535	SEQ ID NO.984	SEQ ID NO.985	X57337	SEQ ID NO.1025	SEQ ID NO.1026
D85904	SEQ ID NO.986	SEQ ID NO.987	X83972	SEQ ID NO.1027	–
AA071853	SEQ ID NO.988	–	X58251	SEQ ID NO.1028	SEQ ID NO.1029
U47543	SEQ ID NO.989	SEQ ID NO.990	AA017867	SEQ ID NO.1030	–
W13646	SEQ ID NO.991	–	W97817	SEQ ID NO.1031	–
U58513	SEQ ID NO.992	SEQ ID NO.993	W75007	SEQ ID NO.1032	–
U70210	SEQ ID NO.994	–	U27838	SEQ ID NO.1033	SEQ ID NO.1034
U18372	SEQ ID NO.995	–	U16163	SEQ ID NO.1035	SEQ ID NO.1036
AA153484	SEQ ID NO.996	–	U20264	SEQ ID NO.1037	–
AA166440	SEQ ID NO.997	–	W18950 W16377	SEQ ID NO.1065	–
W08897	SEQ ID NO.998	–		SEQ ID NO.1066	–
AA162708	SEQ ID NO.999	–			

All the publications, patents and patent applications cited in the present specification are incorporated herein by reference in their entireties.

INDUSTRIAL APPLICABILITY

According to the present invention, ischemic conditions are examined by analyzing

gene expression in a test tissue. Application of such analysis to prevention and treatment of ischemia is expected. There is also provided a novel method of screening for prophylactics and therapeutics for ischemia using, as an indicator, the expression levels of genes which are expressed specifically under ischemic conditions.

SEQUENCE LISTING FREE TEXT

SEQ ID NO: 1067: synthetic DNA

SEQ ID NO: 1068: synthetic DNA

CLAIMS

1. A method for examining ischemic conditions, comprising measuring the expression levels of particular genes in a test sample or determining the expression profile of a gene group in the sample comprising a plurality of genes selected from said particular genes.

2. The method according to claim 1, wherein said particular genes are:

(a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or

(b) genes functionally equal to the genes having any of said nucleotide sequences or genes functionally equal to the genes encoding any of said amino acid sequences.

3. The method according to claim 1 or 2, wherein the measurement of the expression levels or the determination of the expression profile is carried out with a DNA chip.

4. The method according to claim 3, wherein the DNA chip is a synthetic-type DNA chip.

5. The method according to any one of claims 1 to 4, wherein the ischemic condition is at least one selected from the group consisting of compressive ischemia, occlusive ischemia and vasospastic ischemia.

6. A DNA chip for examining ischemic conditions, carrying a part or all of the following genes (a) or (b) immobilized on its surface:

(a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or

(b) genes functionally equal to the genes having any of said nucleotide sequences or genes functionally equal to the genes encoding any of said amino acid sequences.

7. The DNA chip according to claim 6, wherein the ischemic condition is at least one selected from the group consisting of compressive ischemia, occlusive ischemia and vasospastic ischemia.

8. A method of screening for ischemic condition-improving drugs or therapeutics for ischemic diseases, comprising selecting candidate drugs using as an indicator whether or not:

(a) the expression levels of particular genes of which expression levels change under ischemic conditions return to a normal expression levels; or

(b) the expression profile of a gene group comprising a plurality of said particular genes returns to a normal expression profile;

by the administration of a drug to a test animal or test cell, wherein the returning to the normal expression levels or normal expression profile indicates that said drug is a candidate drug.

9. The method according to claim 8, wherein said particular genes of which expression levels change under ischemic conditions are:

(a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or

(b) genes functionally equal to the genes having any of said nucleotide sequences or genes functionally equal to the genes encoding any of said amino acid sequences.

10. The method according to claim 8 or 9, wherein the ischemic condition is at least one selected from the group consisting of compressive ischemia, occlusive ischemia and vasospastic ischemia.

11. A computer-readable record medium in which the following data (a) or (b) have been recorded:

(a) expression level data of genes of which expression levels change under ischemic conditions; or

(b) expression profile data of a gene group comprising a plurality of genes selected from said genes.

12. The record medium according to claim 11, wherein said genes of which expression levels change under ischemic conditions are:

(a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or

(b) genes functionally equal to the genes having said nucleotide sequences or genes functionally equal to the genes encoding said amino acid sequences.

13. The record medium according to claim 11 or 12, wherein the ischemic condition is at least one selected from the group consisting of compressive ischemia, occlusive ischemia and vasospastic ischemia.

14. A computer-readable record medium in which a program that directs a computer to execute the following procedures has been recorded:

(a) procedures to input expression level data or expression profile data of particular genes in a test sample;

(b) procedures to record the input data;

(c) procedures to check the recorded data with already recorded expression level data or expression profile data of said genes under ischemic conditions;

(d) procedures to determine whether the test sample is under ischemic conditions or not based on the checking results obtained in (c); and

(e) if the test sample has been determined as being under ischemic conditions, procedures to identify the clinical stage of the ischemic conditions of the test sample based on the

checking results obtained in (c).

15. The record medium according to claim 14, wherein said genes are:

(a) genes having any of the nucleotide sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066 or genes encoding any of the amino acid sequences shown in SEQ ID NO: 1 through SEQ ID NO: 1066; or

(b) genes functionally equal to the genes having any of said nucleotide sequences or genes functionally equal to the genes encoding any of said amino acid sequences.

16. The record medium according to claim 14 or 15, wherein the ischemic condition is at least one selected from the group consisting of compressive ischemia, occlusive ischemia and vasospastic ischemia.

Fig. 1

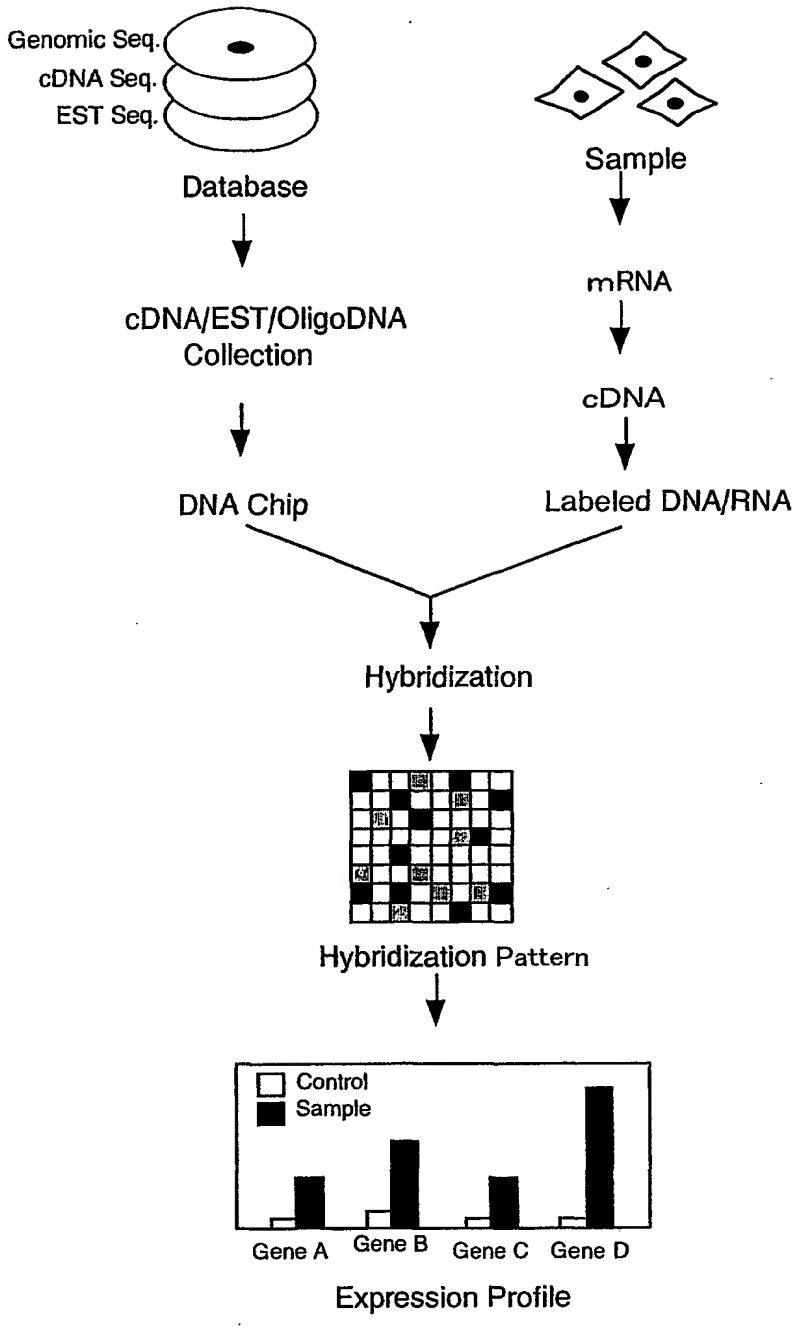

Genomic Seq.
cDNA Seq.
EST Seq.
Database
cDNA/EST/OligoDNA
Collection
DNA Chip
Sample
mRNA
cDNA
Labeled DNA/RNA
Hybridization
Hybridization Pattern
Control
Sample
Gene A
Gene B
Gene C
Gene D
Expression Profile
Expression Level

Fig. 2

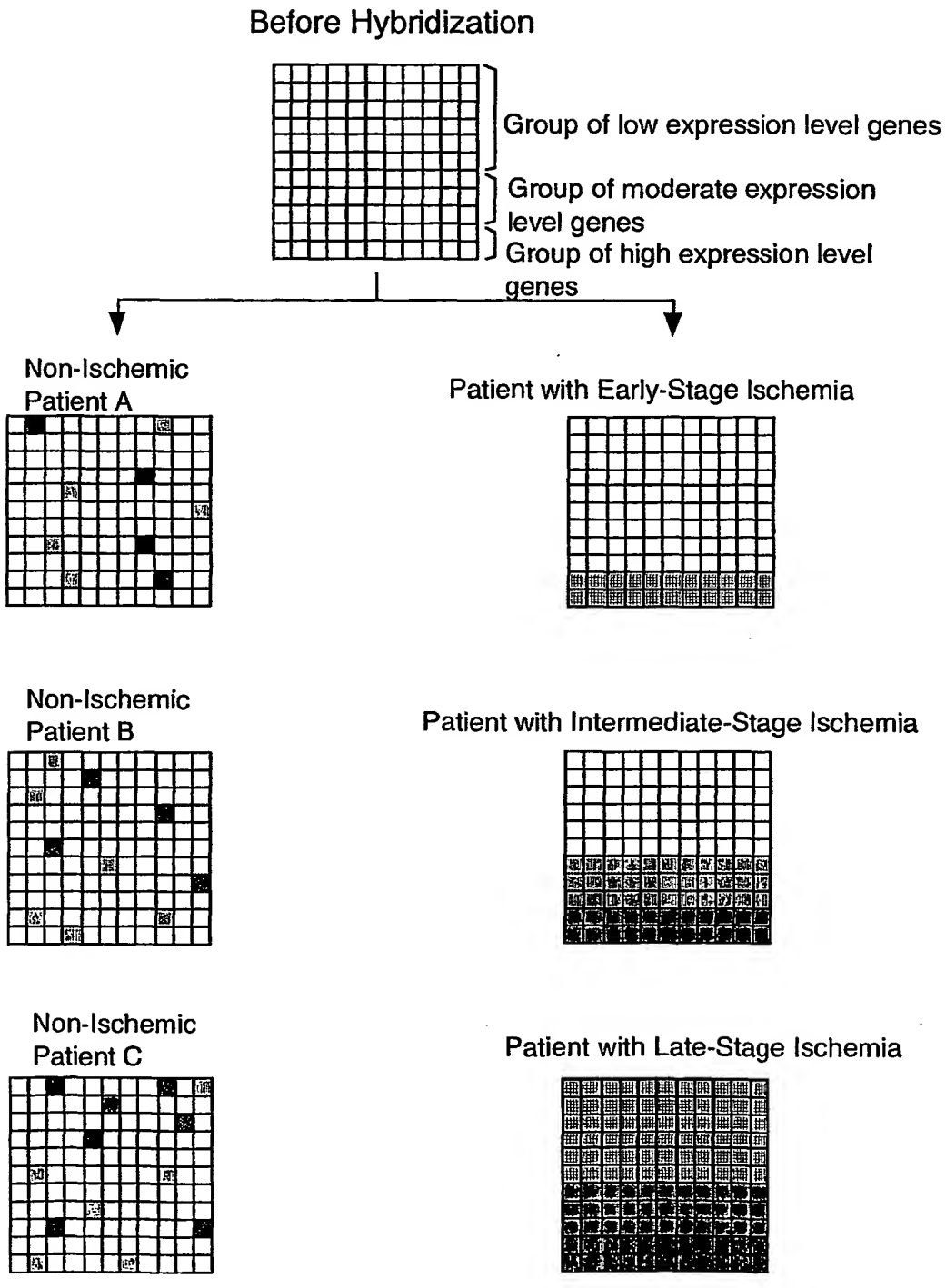

Before Hybridization
Group of low expression level genes
Group of moderate expression level genes
Group of high expression level genes
Non-Ischemic Patient A
Patient with Early-Stage Ischemia
Non-Ischemic Patient B
Patient with Intermediate-Stage Ischemia
Non-Ischemic Patient C
Patient with Late-Stage Ischemia

Fig. 3

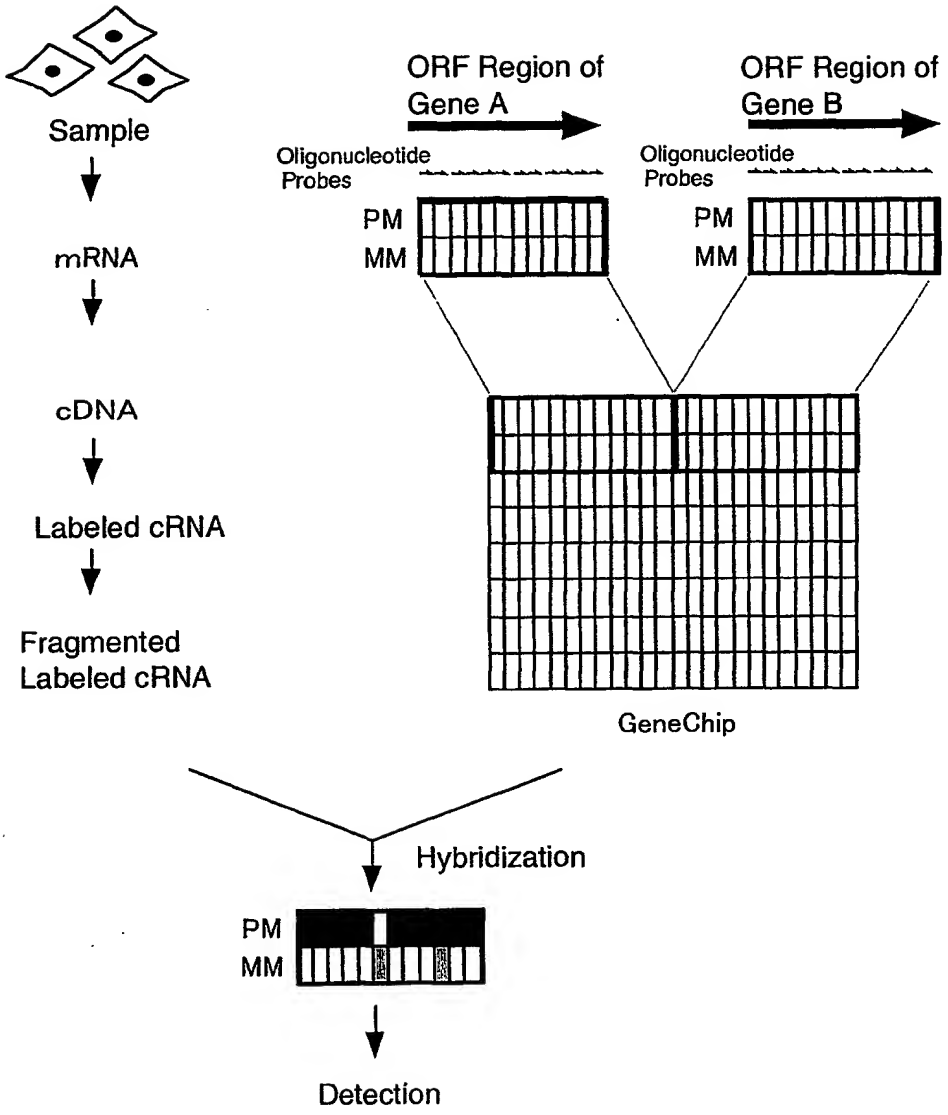

Sample
mRNA
cDNA
Labeled cRNA
Fragmented
Labeled cRNA
ORF Region of
Gene A
ORF Region of
Gene B
Oligonucleotide
Probes
Oligonucleotide
Probes
PM
MM
PM
MM
GeneChip
Hybridization
PM
MM
Detection

Fig.4

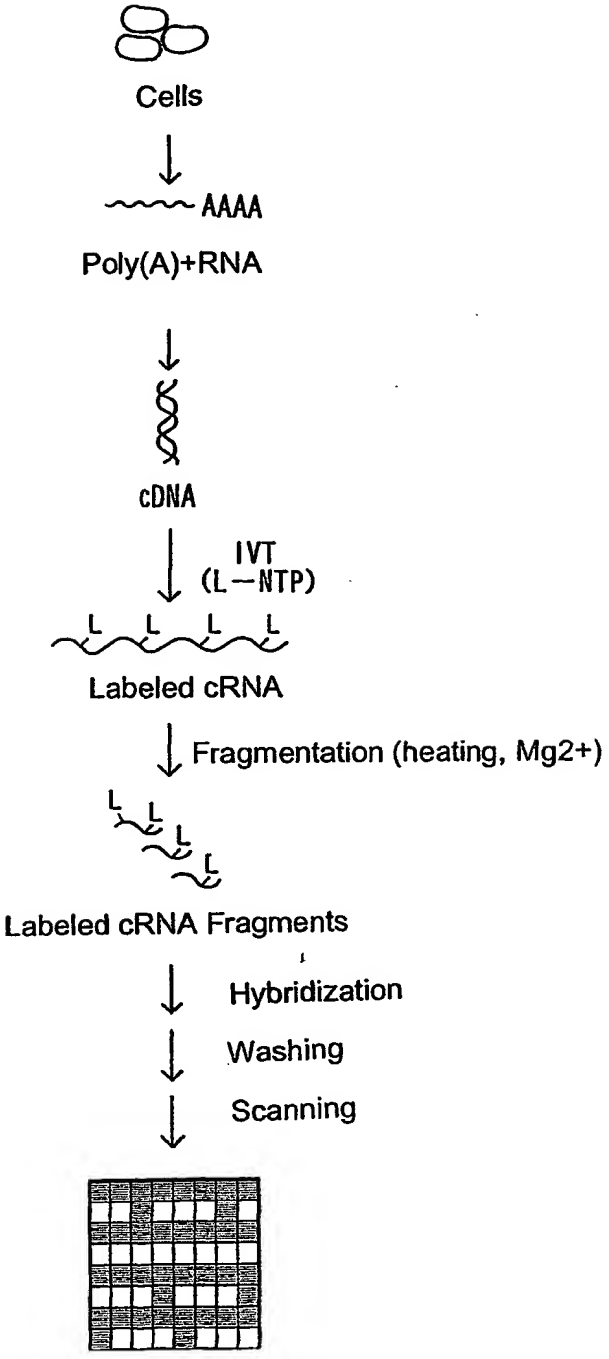

Cells
AAAA
Poly(A)+RNA
cDNA
IVT
(L−NTP)
L L L L
Labeled cRNA
Fragmentation (heating, Mg2+)
L L L L
Labeled cRNA Fragments
Hybridization
Washing
Scanning

Fig.5

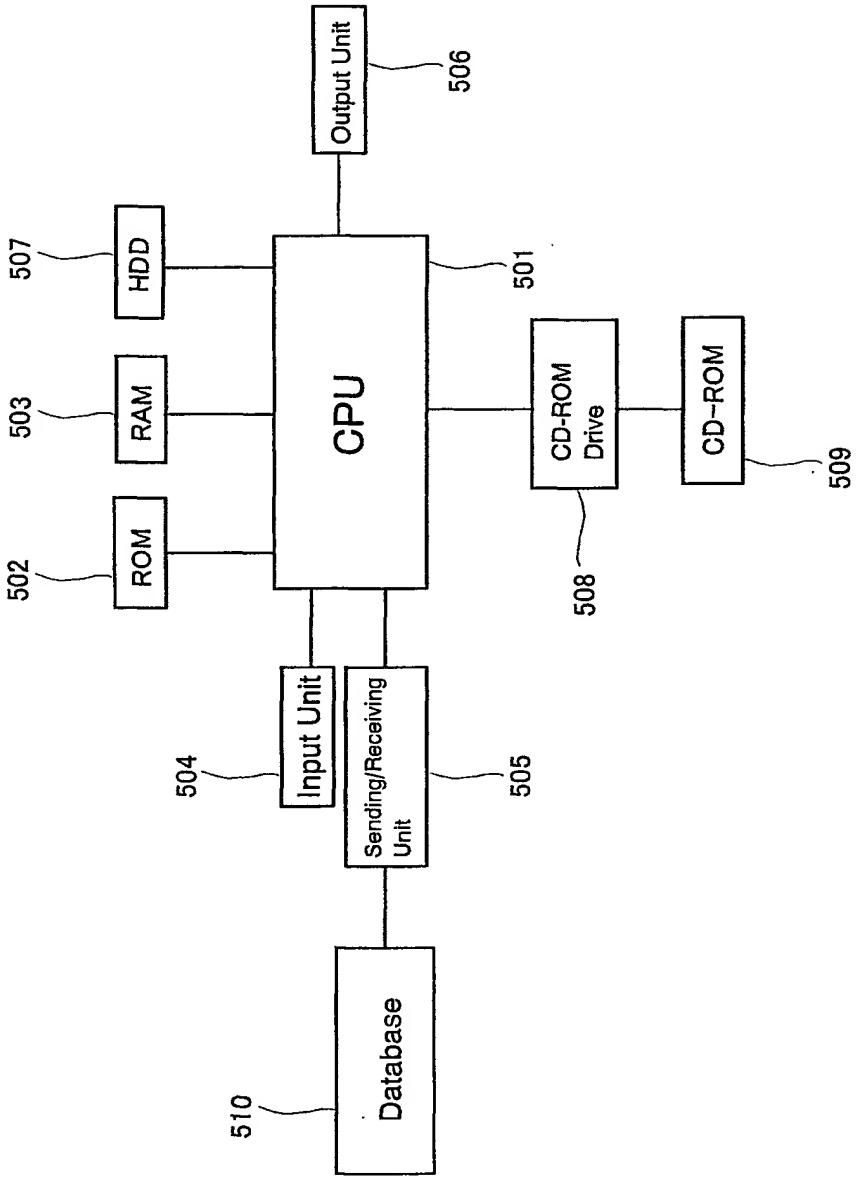
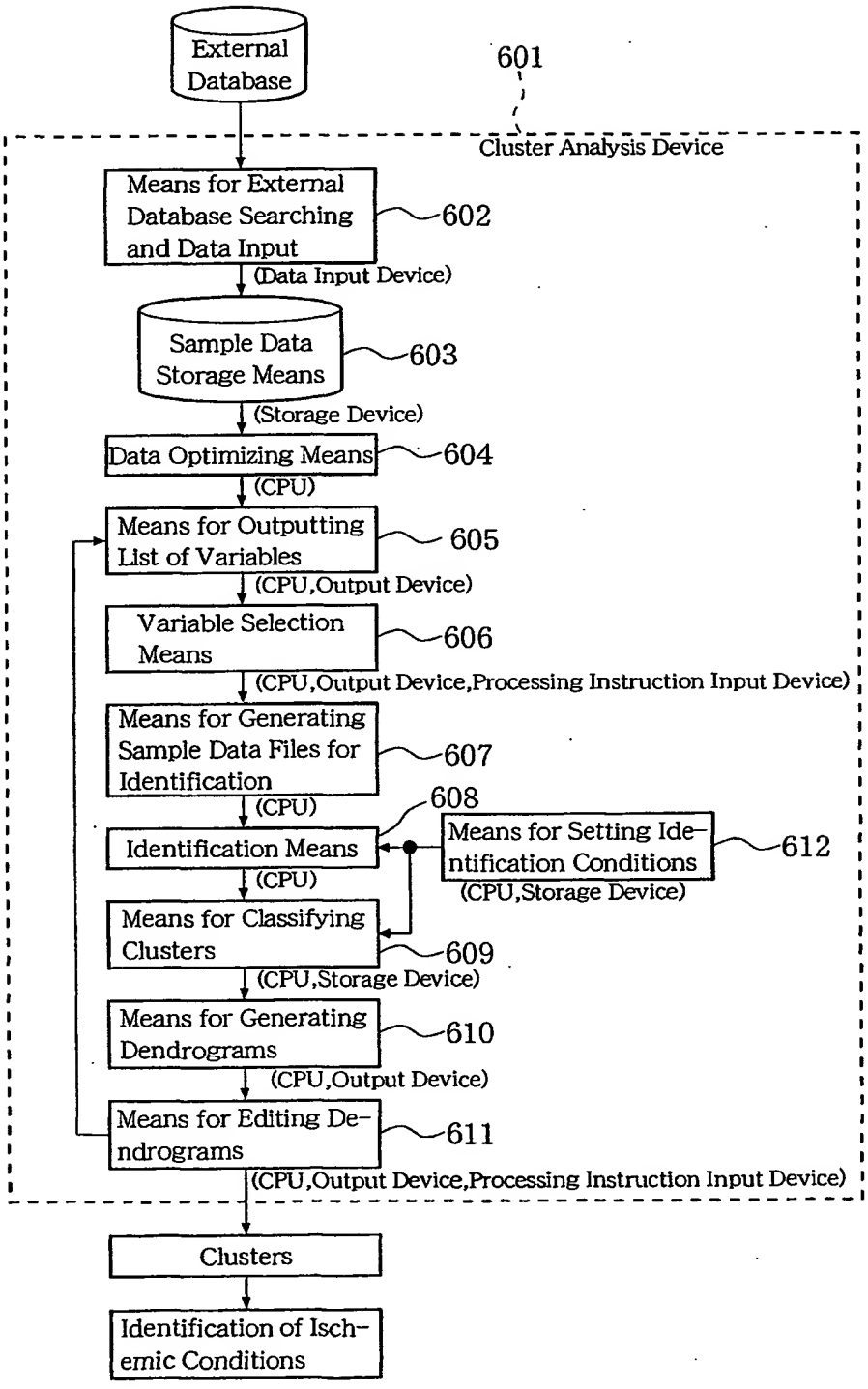

Output Unit
506
HDD
507
CPU
501
RAM
503
CD-ROM Drive
CD-ROM
509
ROM
502
508
Input Unit
504
Sending/Receiving Unit
505
Database
510

Fig.6


External Database
601
Cluster Analysis Device
Means for External Database Searching and Data Input
602
(Data Input Device)
Sample Data Storage Means
603
(Storage Device)
Data Optimizing Means
604
(CPU)
Means for Outputting List of Variables
605
(CPU,Output Device)
Variable Selection Means
606
(CPU,Output Device,Processing Instruction Input Device)
Means for Generating Sample Data Files for Identification
607
(CPU)
608
Identification Means
(CPU)
Means for Setting Identification Conditions
612
(CPU,Storage Device)
Means for Classifying Clusters
609
(CPU,Storage Device)
Means for Generating Dendrograms
610
(CPU,Output Device)
Means for Editing Dendrograms
611
(CPU,Output Device,Processing Instruction Input Device)
Clusters
Identification of Ischemic Conditions

420 425 430

tct aac ctg gtc ttc acc tac tga agggagcgcg ggccggggag aaggtgggcc 2009
Ser Asn Leu Val Phe Thr Tyr
435 440

aagaggcagg agaggagaga gggaaaaaag aaacaaaaaa caaacaaaaa aaaaaaaaga 2069

<210> 461
<211> 440
<212> PRT
<213> Mus musculus

<400> 461
Met Val Gln Gln Thr Asn Asn Ala Glu Asn Thr Glu Ala Leu Leu Ala
1 5 10 15
Gly Glu Ser Ser Asp Ser Gly Ala Gly Leu Glu Leu Gly Ile Ala Ser
20 25 30
Ser Pro Thr Pro Gly Ser Thr Ala Ser Thr Gly Gly Lys Ala Asp Asp
35 40 45
Pro Ser Trp Cys Lys Thr Pro Ser Gly His Ile Lys Arg Pro Met Asn
50 55 60
Ala Phe Met Val Trp Ser Gln Ile Glu Arg Arg Lys Ile Met Glu Gln
65 70 75 80
Ser Pro Asp Met His Asn Ala Glu Ile Ser Lys Arg Leu Gly Lys Arg
85 90 95
Trp Lys Leu Leu Lys Asp Ser Asp Lys Ile Pro Phe Ile Gln Glu Ala
100 105 110
Glu Arg Leu Arg Leu Lys His Met Ala Asp Tyr Pro Asp Tyr Lys Tyr
115 120 125
Arg Pro Arg Lys Lys Val Lys Ser Gly Asn Ala Gly Ala Gly Ser Ala

130 135 140

Ala Thr Ala Lys Pro Gly Glu Lys Gly Asp Lys Val Ala Gly Ser Ser
145 150 155 160

Gly His Ala Gly Ser Ser His Ala Gly Gly Gly Ala Gly Gly Thr Ser
165 170 175

Lys Pro Thr Pro Lys Lys Ser Cys Gly Pro Lys Val Ala Gly Ser Ser
180 185 190

Val Gly Lys Pro His Ala Lys Leu Val Pro Ala Gly Gly Ser Lys Ala
195 200 205

Ala Ala Ser Phe Ser Pro Glu Gln Ala Ala Leu Leu Pro Leu Gly Glu
210 215 220

Pro Thr Ala Val Tyr Lys Val Arg Thr Pro Gln Leu Ala Thr Pro Ala
225 230 235 240

Ala Ser Ser Ser Pro Ser Ser Ala Leu Ala Thr Pro Ala Lys His Pro
245 250 255

Ala Asp Lys Lys Val Lys His Val Tyr Leu Phe Gly Ser Leu Gly Ala
260 265 270

Ser Ala Ser Pro Val Gly Gly Leu Gly Ala Cys Ala Asp Pro Ser Asp
275 280 285

Pro Leu Gly Leu Tyr Glu Asp Gly Gly Pro Gly Cys Ser Pro Asp Gly
290 295 300

Arg Ser Leu Ser Gly Arg Ser Ser Ala Ala Ser Ser Pro Ala Ala Ser
305 310 315 320

Arg Ser Pro Ala Asp His Arg Gly Tyr Ala Ser Leu Arg Ala Ala Ser
325 330 335

Pro Ala Pro Ser Ser Ala Pro Ser His Ala Ser Ser Ser Leu Ser Ser
340 345 350

Ser Ser Ser Ser Ser Ser Gly Ser Ser Ser Ser Asp Asp Glu Phe Glu
355 360 365